Characterization of Ethiopian Durum Wheat Landraces and Cultivars for Processing Quality Using Phenotypic Traits and Simple Sequence Repeat (SSR) Markers

A Thesis Submitted to the Department of Microbial, Cellular and Molecular Biology of Addis Ababa University in partial fulfillment of the requirements for the degree of Master of Science in Biology (Applied Genetics)

By: Temesgen Dagnaw

Advisors:

Dr. Kassahun Tesfaye (PhD)

Dr. Teklehaimanot Haileselassie (PhD)

Dr. Mulatu Geleta (PhD)

September, 2021

Addis Ababa, Ethiopia

Approval letter

Department of Microbial, Cellular and Molecular Biology

College of Natural and Computational Sciences

Addis Ababa University

This is to certify that the thesis prepared by Mr. Temesgen Dagnaw, entitled "Characterization of Ethiopian Durum Wheat Landraces and Cultivars for Processing Quality Using Phenotypic Traits and Simple Sequence Repeat (SSR) Markers" submitted in partial fulfillment of the requirement for the degree of Master of Science in Biology (Applied Genetics) complies to the regulation of the University and meets the accepted standards with respect to originality and quality.

Approved by Examining Board:

Dr. Kassahun Tesfaye Advisor	Signature	Date
Dr. Teklehaimanot Haileselassie Advisor	Signature	Date
Dr. Mulatu Geleta Advisor	Signature	Date
Dr. Mebeaselassie Andargie External examiner	Signature	Date
Dr. Demsachew Guadie Internal examiner	Signature	Date
Dr. Adey Feleke Chair of the Department	Signature	Date

Declaration

I declare that this thesis is my legitimate work and all sources of materials used in this thesis have been appropriately acknowledged. This thesis has been submitted in partial fulfillment of the requirements for the Degree of Master of Science in Biology (Applied Genetics) to the Department of Microbial, Cellular and Molecular Biology of Addis Ababa University and is deposited and available under the rules of the University Library. I solemnly declare that this thesis is not submitted to any other institution anywhere for an award of any academic Degree, Diploma, or Certificate.

Brief quotations from this thesis are allowable without special permission provided that accurate acknowledgment of the source is made. However, permission must be obtained from the author for extended quotation from or reproduction of this thesis in the whole or in part.

Name of the Author: Temesgen Dagnaw

Signature: _____

Place: Addis Ababa University, Addis Ababa.

Date of Submission: September 8, 2021

Abstract

*Durum wheat (*Triticum turgidum *L. ssp.* durum *Desf.) is getting attention in terms of production and area coverage at the global level. Its unique characteristics to produce pasta and related end-products increased the preference for durum wheat by millers and processors. Ethiopia is considered as a center of diversity for durum wheat. However, in Ethiopia, durum wheat is facing serious genetic erosion due to the expansion of teff and bread wheat, and the cultivation of exotic durum wheat materials. The genetic resource of the Ethiopian durum wheat genotypes has been limitedly explored for its processing quality attributes at the level of both phenotypic and molecular diversity. Therefore, the present study was aimed to assess the genetic diversity and evaluate the processing quality attributes of Ethiopian durum wheat genotypes using phenotypic traits and SSR markers. The field experiment was conducted at two locations (Sinana in Bale zone and Chefe Donsa in East Shewa zone) to assess the phenotypic diversity. The phenotypic characterization of genotypes showed a wide range of variability for most quantitative and qualitative traits. The combined analysis of variance (ANOVA) showed highly significant ($P<0.01$) variations among genotypes for the majority of the traits studied. Gluten content (GL) and grain yield (GY) showed high and intermediate heritability, respectively, combined with moderate genetic advance, and grain protein content (GPC) showed intermediate heritability combined with low genetic advance. Both, GPC and GL showed a significant and negative correlation with GY. Cluster analysis based on quantitative traits grouped the genotypes into 5 major clusters. The first four principal components (PCs) explained 64% of the total variation. Based on qualitative traits, high genetic variation was observed among genotypes. Correspondence analysis discriminated cultivars from other populations of landraces. Finally, 104 best-qualified genotypes for processing quality traits were selected and assessed using quality traits associated 14 SSR markers, which had high mean polymorphic information content (PIC) of 0.56 and gene diversity (0.61). High levels of genetic diversity was obtained among all populations ($I = 0.86$; $He = 0.46$). Analysis of molecular variance (AMOVA) revealed high genetic variation within populations (88.35%) and among populations (11.65%). The Neighbor-joining (NJ) clustering and principal coordinate analysis (PCoA) grouped the genotypes into three major clusters. In addition, Bayesian model-based population structure analysis revealed two major genetic groups. Overall, this study revealed high genetic diversity among Ethiopian durum wheat landraces and cultivars, and these genotypes can be used to identify the best genotype/s for processing quality attributes and subsequent use in the Ethiopian durum wheat improvement programs.*

Key Words: durum wheat, Ethiopia, genetic diversity, phenotypic diversity, SSR.

Dedication

I dedicate this thesis to Dr. Kifle Dagne who passed away last year and to all who taught me for being myself.

Acknowledgment

I am very grateful to all individuals and organizations who contributed directly or indirectly to this research. First of all, I would like to address my sincere gratitude to my advisors for allowing me to conduct my research through the SIDA project. All of my advisors deserve special thanks for their interest, advice, and support in every step of this research. I wish to express my deepest appreciation and gratitude to Dr. Kassahun Tesfaye, Dr. Teklehaimanot Haileselassie, Dr. Mulatu Geleta, Dr. Gurja Belay, and Mr. Behailu Mulugeta for believing and inviting me to join the SIDA project which provided me the fund to support this study.

My deepest gratitude and thanks go to Dr. Kassahun Tesfaye for his financial support, constructive suggestions, and morals from the very beginning of the field and molecular experiment to commenting on the write-up. I am very grateful to Dr. Teklehaimanot Haileselassie for his encouragement, guidance, follow-up, and comments started from field trials to the compilation of the write-up. I am grateful to Dr. Mulatu Geleta for the provision of microsatellite primers of durum wheat and PAGE gel red, and for his constructive advice. Mr. Behailu Mulugeta deserves special thanks for his all-rounded support from the beginning of the field and molecular experiment to the end of this study through comments and constructive suggestions on the write-up.

I wish to acknowledge the Department of Cellular, Microbial and Molecular Biology of the College of Natural and Computational Science of Addis Ababa University and to the whole staff members for their kind support and encouragement during the research. I am personally grateful to my colleagues at the plant genetics laboratory especially Dr. Tilahun Mekonen, Mr.Haftom Birhane and Mr. Kiros Tekle for their help and support particularly in identifying and solving the

problems that I encountered in the molecular work, and also Mrs. Hawi Nigussie and Mr. Wubshet Mamo for sharing and facilitating the delivery of resources of the molecular work. I would like to thank the staff and daily laborer of Debrezeit and Sinana agricultural research centers for helping me from the beginning of the trial management to the data compilation.

My family deserves special thanks for encouraging, supporting, loving me in every aspect of my life. All of my teachers, friends, brothers, and sisters who have contributed directly or indirectly to the compilation of the thesis are highly appreciated and deserve special thanks for making my dream true.

Table of Contents

List of Figures

List of Tables

List of Appendices

Acronyms and Abbreviations

AAU	Addis Ababa University
AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of Molecular Variance
ANOVA	Analysis of Variance
CIMMYT	International Maize and Wheat Improvement Center
CTAB	Cetyltrimethyl Ammonium Bromide
DAP	Diammonium Phosphate
DzARC	Debrezeit Agricultural Research Center
EBI	Ethiopian Biodiversity Institute
EIAR	Ethiopian Institute of Agricultural Research
FAO	Food and Agriculture Organization
IGC	International Grain Council
IMARC	International Market Analysis Research Consulting
IPO	International Pasta Organization
ISSR	Inter Simple Sequence Repeat
NJ	Neighbor Joining
NTSYS	Numerical taxonomy and multivariate analysis system
PAGE	Poly-acrylamide Gel Electrophoresis
PCA	Principal Component Analysis
PCoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SARC	Sinana Agricultural Research Center
SNP	Single Nucleotide Polymorphism
SSA	Sub-Saharan Africa
SSR	Simple Sequence Repeat
USDA	United States Department of Agriculture

1. Introduction

The world's modern agriculture is increasingly reliant on wheat. According to the grain market report of International Grain Council (IGC), global wheat production in 2019/20 was 761 million metric tons and the global durum wheat production progressed from 35.4 million metric tons in 2010/11 growing season to 40.7 million tons in 2016/17 season which puts durum wheat as the 10th most important and commonly cultivated cereal crop worldwide. In Ethiopia, durum wheat constituted 10-15 % of the total wheat production in 2018/19 (Brasesco *et al*., 2019).

Durum wheat (*Triticum turgidum* L. ssp. *durum*, Desf.) is an allotetraploid (AABB) species with 28 chromosomes (2n=4x=28), originated through intergeneric-hybridization and polyploidization between *T. uartu* (AA genome) and *Aegilop speltoides* related species (BB genome) (Hancock, 2004). Durum wheat mainly prefers a dry climate, with hot days and cool nights during the growing season and heavy black clay soil type. Ethiopian highlands are major durum wheat producing areas (Hailu Gebre Maryam, 1991; Efrem Bechere *et al.*, 2000). Several studies reported the presence of high phenotypic and genotypic variability in Ethiopian durum wheat landraces. Ethiopia is considered as the center of origin and diversity for tetraploid wheat (Zohary, 1970) and as a center of diversity and secondary center of origin of durum wheat (Vavilov, 1992; Pecetti *et al.*, 1992).

Durum wheat is chosen for making good quality pasta and other processed products because of its kernel hardness, density, high grain protein content, and gluten content and strength. Motalebi *et al.* (2007) mentioned that kernel hardness and vitreousness, its semolina firmness, elasticity and stickiness resistance characteristics make durum to be preferred to produce good quality pasta. In Ethiopia, durum wheat is consumed mainly traditionally and processed by a limited

number of processing sectors mainly for flour, pasta and macaroni. Despite Ethiopia has wide genetic diversity of durum wheat, the production is low both in terms of yield and quality which led to import of pasta and other related products (Brasesco *et al*., 2019). Introduced exotic durum wheat materials to the cultivation program in Ethiopia often failed to adapt to the wide agro-ecological and climatic variations of the country (Dejene Mengistu *et al*., 2015). Hence, characterizing, the genetic diversity of landraces is helpful in recognizing adaptive genotypes for the trait of interest.

Many previous studies used limited numbers of durum wheat genotypes, for example, 87, 70, 16, 49 and 20 Ethiopian durum wheat genotypes were studied by Faris Hailu *et al.* (2006), Yifru Teklu *et al.* (2006), Mohammed Abinsa *et al.* (2012), Dejene Mengistu and Pe (2016) and Daniel Hailegiorgis *et al.* (2017), respectively, and they also used narrow area coverage in terms of landraces collection. Limited researches, for example, 274, 289, 192 and 285 Ethiopian durum wheat genotypes were studied by Dejene Mengistu *et al.* (2015), Dejene Mengistu *et al.* (2016), Admas Alemu *et al.* (2020), and Kefyalew Negisho *et al.* (2021), respectively. In addition, 160 Ethiopian durum wheat genotypes were studied by Meseret Asmamaw *et al*. (2019) and Meseret Asmamaw *et al.* (2020).

Durum wheat is facing genetic erosion by giving way to teff and new bread wheat cultivars in Ethiopia (Bayush Tsegaye and Berg, 2007; Uddin and Boerner, 2008). The production of durum wheat in comparison with bread wheat declined from 60% durum and 40% bread wheat in 1970s to 40% durum and 60% bread wheat in the 1990s (Hailu Gebre-Mariyam, 1991) and then 10-15 % durum and 80-85 % bread wheat in 2018 (Brasesco *et al*., 2019). On the other hand to the erosion, fortunately, Ethiopia is producing 90% of the durum wheat cultivated in SSA and is the only country capable of producing pasta using locally grown grain, while all other SSA countries

import 483 million euro worth of durum grain for pasta production (Sall *et al.*, 2019). This huge gap in SSA and the high level of genetic diversity in Ethiopian durum wheat genotypes creates an opportunity for Ethiopia to expand the agricultural and commercial scope of durum wheat to satisfy domestic and foreign demand. The high level of genetic variation held by Ethiopian durum wheat landraces is a potential for identifying landraces and/or developing cultivars that are highly suitable for the processing quality and agro-ecologies. Therefore, the present study aimed to characterize the Ethiopian durum wheat landraces and cultivars for their processing quality attributes through phenotypic traits and SSR markers. Finally, the result from this study may help to identify the promising durum wheat genotypes with distinct processing quality attributes for superior quality in pasta-making requirements.

2. Objectives

2.1. General Objective:

The general objective of the present study is to characterize the Ethiopian durum wheat landraces and cultivars for their processing quality attributes through morphological traits and SSR markers.

2.2. Specific Objectives:

- To assess the phenotypic diversity of Ethiopian durum wheat genotypes collected from different geographic origins of Ethiopia;
- To identify promising Ethiopian durum wheat genotypes with distinct processing quality attributes for superior quality in pasta making requirements and;
- To assess the molecular diversity of Ethiopian durum wheat germplasm collected from different geographic origins of Ethiopia.

3. Literature Review

3.1. Origin and Domestication of Durum Wheat

Wheat, under the genus *Triticum*, is a grass widely cultivated for its seed and also considered as the first crop to be domesticated and cultivated by man that began possibly between 18,000-12,000 B.C. in the Middle-East and Abyssinian regions (Faris Hailu, 2011). According to an archeological analysis of Feldman *et al.* (2007), wild emmer (*Triticum dicoccoides*) was first cultivated in Southern Levant around 9600 B.C. Archaeological evidence suggests that naked emmer reached Ethiopia (local name of Aja) approximately 5000 years ago, probably arriving from the Levantine, through Egypt, along the Silk Road (Luo *et al.*, 2007).

Most tetraploid wheats (AABB) are derived from wild emmer through hybridization between a wild grass, *Triticum Uratu* and a wild goat grass *Aegliops speltoides* long time ago via domestication (Hancock, 2004) and derived by natural selection. Then, the human-driven tetraploid wheat evolution process led to the domestication of wild emmer to emmer wheat. The evolution was due to the accumulation of mutations that occurred during the cultivation of wild emmer which formed the basis for selection. Hence, continued evolution under domestication evolved domesticated emmer wheat to other closely related tetraploid wheat species such as *T. turgidum* L. ssp. *durum, T. turgidum* L. ssp. *turgidum, T. turgidum* L. ssp. *polonicum,* and *T. turgidum* L. ssp. *aethiopicum* (Morris and Sears, 1967).

The primary origination of durum wheat was suggested to have occurred probably between 12,000 to 10,000 years ago, in the West Levantine (Hakan, *et al.*, 2010). However, Pecetti *et al.* (1992) stated the distinctiveness of the Ethiopian durum wheat germplasm with unique morphology and without an evident allelic similarity to the primary origin in the Levantine. In

addition, Kabbaj *et al.* (2017) molecular data indicated that Ethiopian farmers had also developed durum wheat anew through further domestication of emmer wheat that gave rise to *T. turgidum* ssp. *aethiopicum*.

3.2. Agro-ecologies of Durum Wheat

Durum wheat mainly prefers a dry climate, with hot days and cool nights during the growing season which is shown in the Mediterranean and temperate climates (Brankovic *et al.,* 2014). Traditionally, durum wheat is of narrower adaptation; it is grown in dry and warm regions where limited water supply and prolonged high temperatures during grain development are obstacles for common wheat cultivation (Rachon and Szumilo, 2009).

Durum wheat, like other tetraploid wheat, is widely grown in heavy black clay soil (vertisols) of Ethiopian highlands in 1500-3000 meters above sea level (m.a.s.l.) altitudinal range and 1900-2700 m.a.s.l. is the most suitable with 600 to 2000 mm range of annual rainfall (Hailu Gebre-Mariyam (1991). Efrem Bechere *et al.* (2000) reported that Shewa, Gojam, Gonder, Wello, Tigray, and the highlands of Harergie, Arsi and Bale as major production areas of durum wheat in Ethiopia.

3.3. Global Durum Wheat Production

Based on IGC, durum wheat is the 10^{th} most important and commonly cultivated cereal crop worldwide with a yearly average production of 40.7 million tons in 2016/17 which represents 5% of total wheat production with a planting area of 16Mha globally. Durum is produced primarily for making pasta and the majority of the global durum wheat is produced and consumed in North America, Europe, and South America, from which Canada and Turkey are among the leading producing countries with an estimated 2.2Mha each (USDA 2015), followed by Algeria, Italy

and India, each cultivating over 1.5 million ha (Sall *et al.*, 2019). France, Greece, Morocco, Pakistan, Portugal, Kazakhstan, Russia, Spain, and Tunisia cultivate durum wheat on between 0.5 and 0.8 million ha annually (USDA 2015). In Africa, 3.6Mha has been cultivated for durum wheat and among this Algeria, Morocco, Tunisia, Ethiopia, and Egypt are the highest producers. However, €4.16 billion worth of durum wheat grain was imported to Africa in 2013 to support the pasta and couscous market (Sall et al., 2019).

3.4. Durum Wheat Production in Ethiopia

Ethiopia is the largest producer of durum wheat in SSA with approximately 0.60Mha, which covered 90 % of the SSA durum wheat production (Sall *et al.*, 2019). Shewa, Gojam, Gonder, Wello, Arsi, Bale and Tigray parts of Ethiopia are the predominant producer of durum wheat (Tesfaye Tesemma *et al.*, 1991). Durum is mainly produced by smallholder farmers from highlands (1800 and 2800 m.a.s.l) which are having a relatively low temperature, high rainfall, and vertisol soils with waterlogging character (Efrem Bechere *et al.*, 1996b). To avoid the waterlogging problem of the soil durum wheat is planted late in the growing season and it continues to grow during the dry period on residual moisture (Efrem Bechere *et al.*, 1996b). According to Sall *et al.* (2019), late planting minimizes the potential yield in favor of the higher protein content. Durum wheat is favored by farmers practicing mixed farming systems, as it is a good straw yielder for livestock feeding (Adugna Tolera *et al.*, 2007).

In Ethiopia, the production of durum wheat in comparison with bread wheat declined from 60% durum wheat and 40% bread wheat in 1970s to 40% durum wheat and 60% bread wheat in 1990s (Hailu Gebre-Mariyam, 1991) and then 10-15 % durum wheat and 80-85 % bread wheat in 2018 (Brasesco *et al.*, 2019). This decline in production indicated farmers' preference to cultivate

durum wheat is decreasing from time to time by giving preference to bread wheat cultivars and teff (Bayush Tsegaye and Berg, 2007; Uddin and Boerner, 2008; Asfaw Negassa *et al.*, 2013).

3.5. Stresses in Durum Wheat Productivity

Durum wheat productivity faces abiotic (frost damage, drought, salinity, etc.) and biotic (weed, disease, insect pests, etc.) stresses that impact yields and end-users quality desired by processors. The major production losses in wheat are caused more by abiotic stresses such as drought, salinity, and high temperature than by biotic insults (Abhinandan et al., 2018). Among the abiotic stresses in wheat growth and development, drought, salinity, and high and low temperatures are the major ones. The timing of drought and salinity stresses is critical as certain developmental stages appear to be more sensitive to osmotic damage (Mickky and Aldesuquy, 2017). According to Terletskaya et al. (2020), the tetraploid species T. dicoccum is recognized as a potential source of drought tolerance in wheat. Temperature is a significant determining factor in agronomy practices starting from planting to harvesting. A meta-analysis of 1,700 published simulations (Challinor et al., 2014) predicted a significant yield loss in wheat in temperate and tropical regions with every 2°C rise.

Among the biotic stresses of durum wheat, there are many diseases caused by different microorganisms from fungi to bacteria and viruses. However, pathogenic fungi raised a significant challenge to durum wheat production globally (Bakala et al., 2021). The major diseases in durum wheat involve stripe rust, leaf rust, stem rust, powdery mildew, loose smut, Fusarium head blight (FHB), and wheat blast (WB) (Bakala et al., 2021). Though fungicidal applications offer control, their use is an added cost to farmers besides being unsafe environmentally. Hence growing resistant cultivars is the most effective and efficient control strategy (Rizwan *et al*., 2007). New pathogenic races of these diseases are usually developed

through mutation and most of the diseases are more virulent on durum wheat than hexaploid wheat (Singh et al., 2004) as many of the resistance genes present in the hexaploid wheat (Todorovska et al., 2009). Breeding programs in durum wheat are targeting to develop cultivars that are stable to biotic and abiotic stresses and improved for yield and quality traits (Ellis et al., 2000). Hence, the development of such durum wheat cultivar is highly dependent on the variability of targeted traits, identification, selection, and transfer of superior genes. (Ellis et al., 2000).

3.6. Genetic Diversity of Durum Wheat in Ethiopia

Falconer and Mackay (1996) defined genetic diversity as the occurrence of variability among individuals due to differences in genetic composition. It is also defined as the variety of alleles and genotypes present in a population and this is reflected in morphological, molecular, physiological, and behavioral differences between individuals and populations (Farooq and Azam, 2002). The primary goal of genetic diversity study is identifying an allele or alleles that give the ability to survive in changing and/or stressful habitats. The analysis of genetic diversity can be a useful tool to get information about the genetic diversity of the landraces and varieties that may change the direction of breeding programs (Khlestkina *et al.,* 2004).

Knowledge on genetic diversity is helpful to estimate the loss of genetic diversity, conservation, improvement, evolutionary studies, and so on. The greater the genetic diversity within a species implies the greater that species have a chance of long-term survival and flourishing (Frankel *et al.,* 1995). Genetic diversity is used as a source of novel traits and alleles for plant breeding although it's challenging, particularly to face the unpredictable changing climates and new end-user demands (Tester and Langridge, 2010). Assessing the level and structure of genetic diversity in crops may be a prerequisite for plant breeding and conservation programs. It is

generally thought that persistent choice of crosses among genetically related cultivars results in a narrowing of the genetic base of the crops on which the present-day agriculture is based (Gepts, 2006). This genetic uniformity of crops has driven a few destructive attacks of pests and diseases (van de Wouw *et al.*, 2010).

Ethiopia is very rich in terms of crop diversity, being a primary or secondary center of origin for several crops, including sorghum, durum wheat, barley, and coffee (Brasesco *et al*., 2019). Among tetraploids wheat of Ethiopia studied by Faris Hailu *et al.* (2005), *T. durum* showed the most diversity within species followed by *T. turgidum, T. aethiopicum* and *T. dicoccon*. Many studies reported the existence of high phenotypic and genotypic variability in Ethiopian Durum wheat genotypes (Yifru Teklu and Hammer, 2006; Getachew Belay *et al.*, 1993 and Pecetti *et al.*, 1992). For example, the violet grain and beardless or half-bearded hard durum wheat shows high variation among different parts of Ethiopia (Yifru Teklu and Hammer, 2006; Getachew Belay *et al.*, 1993; Tesfaye Tesemma *et al.*, 1991; Jain *et al.*, 1975).

For durum wheat, Zohary (1970) considered Ethiopia as the center of origin although the diploid einkorn wheat and the hexaploid wheat do not seem to be native to the Ethiopian gene center, whereas Vavilov (1992) and Pecetti *et al.* (1992) considered Ethiopia as a center of diversity. In Ethiopia, the Ethiopian Biodiversity Institute (EBI) has established a holding of over 7000 accessions collected from different parts of Ethiopia (Dejene Mengistu *et al*., 2015). The diverse agro-ecologies, farming system and selection criteria can be counted as factors contributing to the existence of adequate genetic diversity in Ethiopian durum wheat landraces and adaptation to the changing environment (Dejene Mengistu *et al*., 2015). The high level of genetic diversity held by Ethiopian durum wheat is a powerhouse and potential for identifying landraces or developing cultivars that are highly suitable for quality product in the processing sector. Mulatu

Aberra and Tilahun Bayisa (2017) reported the registration of newly released durum wheat variety entitled 'Bullallaa' which showed high protein content (13.3%) and 30.3% gluten content with high thousand grains weight (45.1g), test weight (83.0 kg/L) and good color, that exhibit traits preferred for pasta and related products.

Several previous studies revealed that Ethiopian tetraploid wheats are facing the problem of genetic erosion (Yifru Teklu and Hammer, 2006; Bayush Tsegaye and Berg, 2007; Sall *et al.*, 2019). In Ethiopia, of the overall genetic erosion of tetraploid wheat, 32 % erosion was found for *T. durum* (Yifru Teklu and Hammer, 2006) and the most important factors for this erosion were reduction in cultivated area due to expansion of modern released cultivars of hexaploid wheat and teff (Yifru Teklu and Hammer, 2006; Bayush Tsegaye and Berg, 2007; Sall *et al.*, 2019).

3.7. Global Durum Wheat Processing

The very large amount of genetic diversity contained by durum wheat has widened the importance of this crop from traditional consumption to advanced food processing industries. Now a day's durum wheat is cultivated in developed countries largely to support the food processing industries mainly to produce pasta and macaroni. According to International Pasta Organization (IPO), the annual production of pasta was 16.5 million tons in 2019 and IMARC reported the global pasta market reached a value of $21.8 billion which showed a high increment due to the Covid-19. Europe, South America, and the USA are the leading producers and consumers of pasta in the world, and Africa accounts only for 14.1% of total pasta production, mainly in Tunisia, Egypt and South Africa (IPO, 2021).

Although the yield potential of durum wheat is about 10% lower than bread wheat (FAO, 2019), durum wheat has remained a staple food for smallholder farmers due to adaptation to climatic

stresses (Mulugeta Negassa, 1986; Rachon and Szumilo, 2009), 20 to 40% higher price than bread wheat, millet, maize and sorghum (Sissons, 2004), and good straw yielder for livestock feeding (Adugna Tolera *et al.*, 2007). In addition to these benefits, the existence of a growing value chain of pasta in the global market rescued the crop from extinction and remained under cultivation (Sall *et al.*, 2019).

3.8. Durum Wheat Processing in Ethiopia

Durum wheat is one of the major industrial crops in Ethiopia. Technologies at different scales are used in durum wheat processing, ranging from small-scale (millers) to large-scale (flour factories and processors). In Ethiopia, most processors are in the East Shewa zone (Adama, Bishoftu, Dukem, and Addis Ababa) and Southern Ethiopia (Shashamene and Hawassa). Processors in Ethiopia, medium-to-large, rely on government-subsidized imported durum wheat as a raw material for milling than locally produced because of its supply decrement after Meher (short-rain) season, lack of uniform quality, and expensive price (World Bank, 2018). The pasta production covers about 20-30 % of the production of the country as a whole (150,000 tons) (Brasesco *et al*., 2019). However, according to the Italian Agency for Development Cooperation, such production does not meet the pasta demand of the country as 30,000 tons of pasta was imported in 2014 (Brasesco *et al*., 2019).

As Ethiopia is producing 90% of the durum cultivated in SSA, it is the only country capable of producing pasta using locally grown grain, while all other SSA countries import 483 million euros worth of durum grain in 2013 from Canada, Turkey, and the USA for pasta production requirements (Sall *et al.*, 2019). Hence, this huge gap creates a golden opportunity for Ethiopia to increase the agricultural and commercial scope of durum wheat to satisfy the domestic and SSA demand.

3.9. Processing Quality in Durum Wheat

Quality is an important aspect of durum wheat. High-quality pasta begins with good quality wheat (Mariani *et al.,* 1995). Durum means "hard" in Latin, and *T. durum* is the hardest of all wheat. Because of its kernel hardness, density, high protein content, high gluten content and strength, vitreous endosperm, which gives the semolina good firmness, elasticity, surface disintegration and stickiness resistance (Motalebi *et al.,* 2007), and unique amber endosperm color, which gives the pasta the golden color (Sissons, 2016), durum wheat is chosen for producing premium pasta and related products. The quality of durum wheat affects the quality of pasta from its color up to its taste (Sissons, 2016). The major determinants of wheat quality are the endosperm texture (physical properties that affect milling, semolina and processing and the protein content (chemical properties that affect protein quantity and quality, and pigment content) (Pasha *et al*, 2010).

3.9.1. Physical Qualities in Durum Wheat

Assessment of the potential to make good pasta begins with the grain quality. Such grain quality includes visual appearance, test weight, thousand grains weight, physical defects, vitreousness, moisture content, weather damage, and grain protein percentage (Sissons, 2004). The physical quality of grain is an important grading factor. Depending on the percent of hard and vitreous kernels of amber color, durum wheat is classified into three sub-classes: hard amber durum (75% or more), amber durum (65-75%) and durum (less than 65%). Then each subclass is divided from grade 1 up to 5 based upon test weight per bushel, defects and foreign materials in one kilogram (USDA, 2015).

3.9.1.1. Kernel Vitreousness in Durum Wheat

Kernel vitreousness is a measure of the kernel protein level, color and semolina granulation. It is expressed as the ratio of the weight of vitreous kernels to total sample weight and predicts semolina quality and yield. Vitreous kernels are hard, glassy and translucent whereas non-vitreousness kernels are soft, starchy and optically dense without moisture content effect (Neethirajan *et al.,* 2006). Virtuosity is positively correlated with gluten and protein content (Bilgin *et al.,* 2010; Fu *et al.,* 2017; Taneva *et al.,* 2019a) and semolina color (Bilgin *et al.,* 2010). Vitreous kernels have higher gluten content than starchy kernels (Dexter *et al.,* 1988). But, protein content and/or gluten content are generally considered a more reliable index to predict pasta quality than the degree of vitreousness.

Vitreousness is mainly influenced by environmental conditions such as temperature, light intensity, water availability, and nitrogen fertilization, during grain development rather than genetic factors (Taneva *et al.,* 2019b).

3.9.1.2. Thousand Grains Weight in Durum Wheat

Thousand grains weight (TGW) is the measurement of the weight in grams of 1000 seeds which is linked with test weight and semolina yield. Large kernels would have a higher ratio of endosperm to bran which yields higher semolina than small kernels. In durum wheat, Dejene Mengistu *et al.* (2016) found high heritability (87%) with high genetic advance (44.3) for TGW that indicates environment has less effect on the expression of TGW.

3.9.2. Chemical Properties in Durum Wheat

3.9.2.1. Protein Content in Durum Wheat

Wheat quality is determined by several chemical properties, and the major factor is the protein called gluten, which is roughly 78–85% of total wheat endosperm protein (Weegels *et al.* 1996). The protein content is the most important aspect of durum wheat in determining pasta quality. Protein quantity is influenced by environmental factors, while the quality is genetically determined (Pasha *et al.*, 2010). Protein content affects test weight, TGW, and vitreousness (Mohammed Abinsa *et al.*, 2012), and semolina particle size distribution (Sacchetti *et al.*, 2011). For pasta quality, dough properties are important and its storage protein of wheat endosperm determines dough elasticity, stability and strength. In some countries, protein content is used as a price determinant for farmers (Sissons, 2004). High protein content in durum wheat would give uniform semolina particle size (Sacchetti *et al.*, 2011), which will hydrate evenly when mixed with water and will give strong and elastic pasta. Protein content in durum wheat ranges from 9-18 % (Feillet, 1988), 11-16 % (Turnbull, 2001), 10.7-16.8 % (Sayaslan *et al.*, 2012), and greater than 13% is the preferred amount for high-quality pasta products (Kadkol and Sissons, 2016).

The protein content is influenced mainly by environmental than genotypic factor. The selection of higher-protein cultivars and the proper application of nitrogen fertilizers are helpful for improving the protein content of durum wheat (Pitz, 1992). Wheat storage proteins are composed of non-gluten and gluten proteins. Non-gluten proteins include soluble proteins (albumins, globulins) which constitute 15-20 % of wheat grain protein (Belderok *et al*., 2000) and many of these proteins are enzymes involved in metabolic activity (Zilic *et al.*, 2011). Gluten proteins (glutenins and gliadins) constitute up to 75-85% of wheat grain protein (Belderok *et al*., 2000). Gluten content in grains of durum wheat is the determinant of the gluten content in the flour, that

gluten content of the grain is crucial in processing or pasta making because gluten confer elasticity and viscosity of the dough (Padalino *et al.*, 2019). Despite its importance in the processing sector, eating gluten may confer a disease called Coeliac disease which is an immune-mediated enteropathy in consumers with food-related disorders.

3.10. Factors to Determine Effective Breeding in Durum Wheat

The modern breeding strategy for genetic improvement of durum wheat genotypes, considering the global climate change, is focused on creating high-yield and high-quality cultivars adapted to stress and with good stability. Such efforts can lead to advances in agronomic production and in the food processing industry by increasing value and improving the health properties of final products (Zilic *et al.*, 2011). To achieve this, genetic variability, heritability and genetic advance are the most important prerequisites for designing breeding programs and offer opportunities to plant breeders for selecting highly qualifying genotypes or to combine or transfer genes having desirable traits. In addition, the correlation of different traits to one another is important in plant breeding to make the decision of direct and indirect selection.

3.10.1. Genetic Variability

The very important factor that influences selection in plant breeding is genetic variability. Genetic variability is the tendency of genetic characteristics to vary. Genetic variability can be assessed using several approaches, from the actual DNA sequence to quantitative and qualitative morphological traits. Among the parameters, heterozygosity, number of alleles per locus, percent of polymorphic loci, and allelic and genotypic frequencies are the most widely used measures of variation exists within populations. The presence of genetic variability in crops is essential by providing options for the breeders to develop new varieties and hybrids. Otherwise, crop breeding is impossible in the absence of genetic variability for a particular trait. The main

sources of genetic variations are mutation, recombination, migration, hybridization, segregation, and natural selection in natural populations although for crops human activity via artificial selection has played a major role (Bhandari *et al.*, 2017).

Levels and patterns of genetic variation can be determined by the biological characteristics of the species, and population mating system and structure (Falconer and Mackay, 1996). Comparative variability of traits is evaluated by estimating the genotypic coefficient of variation (GCV) and the phenotypic coefficient of variation (PCV) (Ahmad *et al.*, 2011). The GCV expresses the heritable portion (additive or fixed genes), while the PCV expresses the effects of both genetic and environment on a certain trait (Bello *et al.*, 2007).

3.10.2. Heritability

Heritability is a key parameter in quantitative genetics that determines the response to selection (Johnson *et al*., 1955; Yadav *et al*, 2015). Depending on the number of variances used as a numerator in the calculation, heritability can be grouped into broad-sense and narrow-sense heritability (Hancock, 2004). Heritability in the broad sense is the ratio of the total genetic variance to the total phenotypic variance and in the narrow sense is the ratio of the additive genetic variation to the phenotypic variation (Hancock, 2004). Lourenco *et al.* (2017) expressed heritability as "the extent to which a phenotype is genetically determined." A phenotype is the composite of observable traits of an organism that results from the expression of the genotype of the organism, the influence of environmental factors, and the interactions between both. Thus, heritability investigates the relationship between phenotypic values with phenotypic variance (σ_p^2) and their respective underlying true genotypic values with genotypic variance (σ_g^2).

Heritability is estimated from the degree of resemblance between relatives and its value ranges from zero (when all variation is caused by the environment) to 100% (when all variation is caused by genetic composition) of individual population and is classified as high heritability (>50 %), intermediate heritability (40-50 %) and low heritability (<40 %) according to Singh (1990). The magnitude of heritability indicates the effectiveness with which selection of desirable genotypes and helps to predict the behavior of successive generations. Broad sense heritability only shows if there is sufficient genetic variation in a population and how they will respond to selection pressure (Milatovic, 2010). Heritability values vary with the nature of the test materials and environment (Habtamu Ayalew *et al.*, 2011). The high value of heritability identifies the given trait as highly heritable to a generation which makes the selection procedures simple for breeders. However, it has been accentuated that heritability alone has no practical importance without genetic advance (Najeeb *et al.*, 2009).

3.10.3. Genetic Advance

According to Allard (1960), the improvement in the mean genotypic value of the selected plants over the base population measures the expected genetic gain from selecting the best performing genotypes for a given character. Johnson *et al*. (1955) classified genetic advance as percentage of mean as low (0-10%), medium (10 - 20%), and high (20% and above). The success of genetic advance depends on genetic variability, heritability, and selection intensity (Allard, 1960). High heritability coupled with high genetic advance creates the most suitable condition for selection (Larik *et al.*, 2000). The knowledge on heritability and genetic advance enables to make wise parental selection (Tuhina-Khatun *et al*, 2007) which helps to estimates the phenotype of succeeding generations and the magnitude of genetic improvement through selection (Yadav *et al.*, 2015).

3.10.4. Correlation Studies among Traits

The study of correlation plays a significant role in plant breeding because it helps to determine the magnitude and direction of the relationship between several traits. Simple linear correlation deals with estimate and test of significance of simple linear correlation coefficient (r), which measures the extent of linear association between variables, without classifying one as the cause (independent) and the other as the consequence (dependent) (Gomez and Gomez, 1984). Correlations in quantitative genetics can be grouped into phenotypic correlations (r_P), genetic correlation (r_A) and environmental correlations (r_E) (Falconer and Mackay, 1996). Falconer and Mackay (1996) states the genetic relationship of traits can be occurred as epistatic, pleiotropic, linkage equilibrium, and/or linkage disequilibrium between different genes or due to environmental effects.

Correlation studies may also help to find the best combination of traits to give higher quality and/or yield. Quality-related traits like grain protein content showed a significant positive association with wet gluten content, dry gluten content and TGW and some yield-related traits like test weight showed significant association with yellow pigment content (Mohammed Abinsa *et al.*, 2012). Correlations of characters with yield are useful to identify criteria for indirect selection, to provide reliable information on the nature, extent and direction of selection (Falconer and Mackay, 1996). In crop plants, yield is usually dependent upon the action and interaction of a number of important characters which are generally interrelated and its value can be either negative or positive. For example, in wheat, GY is positively correlated with TGW and negatively correlated with PLH (Khan *et al.*, 2013).

3.11. Marker Systems in Crop Genetic Diversity Study

Different scholars have used several types of markers to investigate the genetic diversity of plant species. Of these: morphological, cytological, and biochemical markers are considered as old conventional markers, whilst molecular markers are considered as the modern tools to assess the genetic diversity within and between crop populations and/or species (Nadeem *et al.*, 2018).

3.11.1. Morphological Markers

Morphological marker systems are the earliest, simple, and inexpensive genetic markers based on phenotypic appearance. Morphological traits such as days to heading and maturity, grain filling period, thousand grains weight, grain yield per plot, plant height, number of effective tillers, number of spikelets per spike, and *etc*. can be used to study the variation among durum wheat genotypes. Previous studies have dealt with the variability of Ethiopian durum wheat landraces for agro-morphological and quality traits (Pecetti *et al.,* 1992; Getachew Belay *et al.,* 1993; Yifru Teklu and Hammer, 2006; Mohammed Abinsa *et al.*, 2012).

3.11.2. Molecular Markers

Molecular genetics made it possible to study genetic variation in the natural population at the DNA level. DNA markers reveal genetic variation among genotypes at the DNA level which is more direct, reliable, precise and efficient for germplasm characterization. DNA-based markers are grouped into two types first non-PCR based (RFLP) and the second is PCR based markers (RAPD, AFLP, SSR, SNP, etc.) (Kumar *et al*., 2009).

3.11.2.1. Simple Sequence Repeats (SSRs)

Simple sequence repeats or microsatellites are short (100-500 bps) tandem repeats of 1 to 6 nucleotide repeats and assumed to be randomly distributed throughout the coding and non-

coding regions of nuclear, mitochondrial DNA and chloroplast DNA (Rajendrakumar *et al*., 2007). Primers (18-25 bp) are designed to known conserved regions flanking the variable microsatellite. SSR alleles are typically co-dominant and their polymorphisms can be detected by PCR and its products can be separated on agarose, polyacrylamide, and/or automated capillary gels (Mun *et al.*, 2006). SSRs detect variation in length that results from changes in the number of repeats or large jumps due to mutation in flanking regions.

Microsatellite markers are the best from other molecular markers (AFLP, RAPD, RFLP, etc.) for genetic diversity studies and fingerprinting of the crop because they have high polymorphism index (Mun *et al.*, 2006), uniformity across the genome, co-dominance (detect polymorphism in total DNA extracts), rapid and easy exchange of data between primers, easy PCR-based assay (Roder *et al.*, 1998), abundant, high reproducibility and no radioactive labeling (Kumar *et al.*, 2009). Despite their benefits, SSR are time-consuming, expensive, species-specific primers, flanking sequence information is needed (Kumar *et al.*, 2009). SSR markers have been applied for characterization of wheat in many studies: Roder *et al.* (1998), Sentayehu Alamerew *et al.* (2004), Yifru Teklu *et al.* (2006), Jemanesh Kifetew *et al.* (2013), Kumar *et al.* (2015), Meseret Asmamaw (2019), and Kumari *et al.* (2019).

4. Materials and Methods

4.1. Experiment I. Phenotypic Diversity

4.1.1. Plant Materials

A total of 420 durum wheat, 386 landraces and 34 cultivars were evaluated. Among this, all the landraces were obtained from Ethiopian Biodiversity Institute (EBI) that were collected from different geographic regions of Ethiopia, and the cultivars were collected from Debrezeit Agricultural Research Center (DzARC) (24) and SARC (10) (Appendix 1). Hereafter, for the sake of simplicity, the term genotypes will be used to represent landraces and cultivars. The geographical locations of the landraces are shown in Figure 1. Landraces collected from nearby geographical zones were considered as an individual population to proportionate the number of samples per populations.

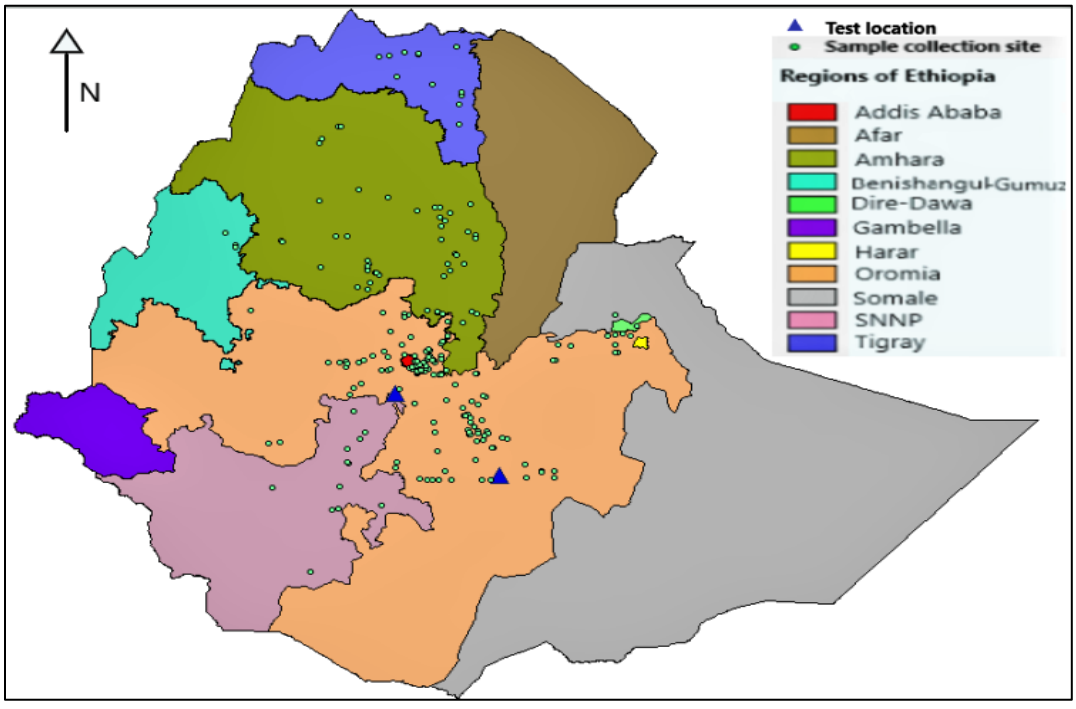


Figure 1. Map of Ethiopia showing the collection site of landraces used in this study.

4.1.2. Test Locations

The field experiments were conducted at two locations (Chafe Donsa and Sinana) in the 2019/20 cropping season. Geographical, metrological and soil-related information of the two locations is presented in Table 1.

Table 1. Geographic, weather and soil description of test locations.

Parameter		Sinana	Chafe-Donsa
Latitude		07º 07’ N	08° 44’ N
Longitude		40º 10’ E	38° 58’ E
Altitude (m.a.s.l)		2400	2450
Mean annual rain fall (mm)		812	851
Rainfall pattern		Bi-modal	Uni-modal
Temperature	Min.	9 ºC	8.9 °C
	Max.	20 ºC	28.3 °C
Soil type		Phaeozems and Cambisols with minor Vertisols.	Vertisols with minor Light soils

m.a.s.l is meter above sea level.; Min is minimum; and Max is maximum. Source: Ethiopian Institute of Agricultural Research (EIAR).

4.1.3. Experimental Design and Field Management

The experiment was laid out in the alpha lattice design with two replications. The alpha lattice design was relatively efficient than RCBD (Masood *et al.*, 2018). Genotypes were randomly assigned to plots within each block with a plot size of 1 m long and the spacing between entry, plants, blocks and replications were 40 cm, 10 cm, 1 m and 2 m, respectively. The fertilizer rates of 100 Kg DAP and 50Kg Urea were used. Nitrogen fertilizer was applied at the planting and jointing stage. All necessary management practices were applied to all experimental plots as recommended for wheat.

4.1.4. Data Collection

The following important phenotypic traits were recorded on plot and plant basis (IBPGR, 1985).

Data recorded on a plot basis were:

- Days to heading (DH): - the number of days from planting to a stage when 50% of the plants in a plot have produced spikes;
- Days to maturity (DM): - the number of days from planting to a stage when 75% of the plants in a plot have reached physiological maturity;
- Grain yield (GY): - the weight of total seeds in gram obtained from a plot area at 12% moisture content;
- Thousand grains weight (TGW): - weight of 1000 kernels in gram at 12% moisture content, as the weight of 250 kernels multiplied by four;
- Stand growth (STD): - the percent of the growing seed from the total of planted seeds;
- Lodging (LDG): - susceptibility to lodging is scored before harvest on a scale similar to percentages. Named as 0 no plant lodged, 1 very few lodged, 5 almost 50% lodged, and 9 nearly all plants lodged;
- Vitreousness (VTR): - the extent of glass-like appearance when seeds are transversely sectioned. Visual measurement was used by coding as 1 as soft and starchy, 2 as partly vitreous, and 3 as vitreous;
- Kernel Color (KCL): - visual measurement of the color of the seed obtained from a plot grouped as 1 for white, 2 for brown, and 3 for purple;
- Glume Hairiness (GLH): - hairy like structure on the outer side of sterile glume grouped as 0 for hairless, 1 for less hairy, and 3 for highly hairy;

- Glume Color (GLC): - the color observed on the outer glume of plants by coding 1 as white to yellow, 2 as red to brown, and 3 as purple to black and
- Spike Density (SPD): - visual measurement of the closeness of spikelets on spikes coded as 3 for lax, 5 for intermediate, and 7 for dense.

Data recorded on a plant basis were:

- Plant Height (PLH): - the average height of five plants in cm from the ground level to the tip of the spike at maturity;
- Number of tillers per plant (NET): - the average number of effective tillers per plant;
- Spike length (SPL): - the average length, in cm, from base to tip of a spike excluding awns and
- Number of spikelets per spike (SPS): - the average number of spikelets per spike from five typical spikes selected from a genotype.

Chemical characters include: grain protein content (GPC), gluten content (GL), and moisture content were measured using Near-Infrared Spectroscopy (NIRS) as described in AOAC (2016) Method 997.06 (32.2.03A). Finally, the values of GPC, GL and GY were adjusted to 12% moisture content. NIRS is widely used to quantify the composition of cereals because of its accuracy and reliability, and for being fast, easy-to-use, and inexpensive.

4.1.5. Statistical Analysis

4.1.5.1. Analysis of Variance (ANOVA)

Prior to computing ANOVA, the collected quantitative data were checked for the elimination of error and then normality checked by Shapiro Wilk test using shapiro.test function of Stats package (R Core Team, 2020) in R software. ANOVA is helpful to partition the overall

variations into components. Effect of replication, location, block and all interactions were tested. ANOVA was calculated for each and combined location. To identify whether traits can be combined from the two locations for the ANOVA, homogeneity variance of error was evaluated by Bartlett test using bartlett.test of Stats package (R Core Team, 2020) in R software. ANOVA was done using aov function of Stats package (R Core Team, 2020) in R statistical software based on alpha lattice design linear additive model.

$$P_{ijk} = \mu + \tau_i + \pi_j + b_k + \xi_{ijk}$$

Where: P_{ij} = phenotypic value of i^{th} treatment under j^{th} replication and k^{th} block; μ = grand mean; τ_i = the effect of i^{th} treatments; π_j = the effect of j^{th} replication; b_k = the effect of k^{th} block; and ξ_{ijk} = experimental error.

4.1.5.2. Estimation of Variance Components

Phenotypic, genotypic and environmental variances were computed from the respective Mean Squares (MS) of combined locations as described by Singh and Chaudhary (1977):

Genotypic variance = $\sigma_g^2 = \frac{MS_g - MS_{g\times l}}{rl}$

Genotype x location interaction variance = $\sigma_{g\times l}^2 = \frac{MS_{g\times l} - MS_e}{r}$

Phenotypic variance = $\sigma_p^2 = \sigma_g^2 + \frac{\sigma_{g\times l}^2}{l} + \frac{\sigma_e^2}{rl}$

Where: MS_g = mean square of genotype; MS_{gXl} = mean square due to genotype by location interaction; MS_e = error mean square; r = number of replications and l = number of locations

Broad sense heritability (H_b) was calculated using the formula presented by Allard (1960):

Broad sense heritability = $H_b = (\sigma_g^2 / \sigma_p^2) \times 100$

Where: σ_g^2 and σ_p^2 = genotypic and phenotypic variance, respectively.

The genotypic and phenotypic coefficient of variation, expected genetic advance and genetic advance as percent of mean computed as described by Johnson *et al*. (1955):

Genotypic coefficient of variation = GCV = 100 x σ_g/μ

Phenotypic coefficient of variation = PCV = 100 x σ_p/μ

Genetic advance = GA = k x H_b x σ_p

GA as percent of mean = GAM = (GA/μ) x 100

Where: σ_g and σ_p = genotypic and phenotypic standard deviation, respectively; k = constant value at selection intensity of 5% (k = 2.06); H_b = broad sense heritability; GA = genetic advance and μ = the grand mean of the trait considered. GCV and PCV values were categorized as high (20%), medium (10-20%), and low (<10%) according to Burton and Devane (1953). The magnitude of GAM is categorized as low (0-10%), moderate (10-20%), and high (>20%) according to Falconer and MacKay (1996).

4.1.5.3. Correlation Coefficient (r)

The pearson correlation coefficient between combined quantitative traits were computed using chart.Correlation function of PerformanceAnalytics package (Peterson *et al.,* 2018) of R software.

4.1.5.4. Cluster and Principal Component Analysis

Cluster analysis is helpful to identify the pattern of diversity among geographically classified populations. The values of the combined quantitative traits were scaled to unit variance and zero for cluster analysis and principal component analysis (PCA) using the scale function of Stats package (R Core Team, 2020) in R software. The scaled data were used to compute the Euclidian distance matrix using dist function of Stats package (R Core Team, 2020) in R software. Then, hclust function of the cluster package (Maechler *et al.,* 2020) was used to create the hierarchical

tree in R software. To identify the best method for clustering, the cophenetic distance was computed for most clustering methods and compared with the original distance using the cophenetic function of Stats package (R Core Team, 2020) in R software. Ward.D2 method was used for clustering because of a high cophenetic value and distinctiveness of clusters of the method.

The best number of cluster was estimated by the Bayesian Information Criterion (BIC) method using Mclust function of mclust package (Scrucca *et al.*, 2016) of R software. The dendrogram was constructed using fviz_dend function of factoextra package (Kassambara and Mundt, 2020) in R software. The mean performance of each trait in each cluster was computed using aggregate function, and used to estimate the pairwise distance between clusters using dist function of Stats package (R Core Team, 2020) in R software.

The scaled data of quantitative traits were further analyzed for PCA using prcomp of Stats package (R Core Team, 2020) in R software. The scatterplot of individual genotype on 2D plane and the coordinate value with contribution of each trait was plotted and estimated using fviz_pca_biplot and get_pca_var function of factoextra package (Kassambara and Mundt, 2020) in R software, respectively.

4.1.5.5. Frequency Distribution and Chi-square Test

The qualitative data were transformed into frequency data using prob.table function of Stats package (R Core Team, 2020) in R software. The frequency data of each population obtained from individual test locations and combined locations was subjected to the chi-square test using chisq.test function of Stats package (R Core Team, 2020) in R software.

4.1.5.6. Shannon Diversity Index

Standardized Shannon diversity index was calculated from the frequency data of qualitative traits as mentioned by Firdissa Eticha *et al*. (2005):

$$H' = (\sum_{i=1}^{R} pi(\mathbf{ln}pi)) / \ln(n)$$

Where n is the number of phenotypic classes of the trait and pi is the proportion of i^{th} phenotypic class. H' was estimated for all qualitative traits and all genotypes. The Shannon diversity index was classified as high (H' ≥ 0.60), intermediate (0.40 ≤ H'< 0.60) and low (H' < 0.40) according to Firdissa Eticha *et al*. (2005).

4.1.5.7. Correspondence Analysis

Correspondence analysis (CA) is a multivariate analysis which is conceptually similar to PCA, but correspondence applies to categorical data. Correspondence analysis of qualitative traits was carried out using Factoshiny function of FactomineR package (Husson *et al.*, 2016) in R software.

4.2. Experiment II. Molecular Diversity

4.2.1. Plant Materials and DNA Extraction

Genotypes with preferred phenotypic characters for quality attributes (grains showed white color, vitreous, high GL and GPC with higher TGW and GY) were selected (Appendix 2) and further assessed using microsatellite markers. Sound seeds of selected genotypes were grown for 15-20 days in the greenhouse of AAU and equal amount of bulk leaf samples from each genotype were dried in silica gel containing plastic bags. Then dried leaves were milled using Geno-grinder machine. Genomic DNA was extracted using the Cetyltriethyl Ammonium Bromide (CTAB) method optimized by Oumer Abdie *et al.* (2020). The DNA extraction and subsequent molecular

marker analysis were done at plant genetics laboratory, Addis Ababa University, Addis Ababa, Ethiopia.

4.2.2. Genomic DNA Quality and Quantity Measurement

DNA quality and quantity test were done by using Nano-drop spectrophotometer and gel electrophoresis. Using gel electrophoresis; by running 3 µl of extracted DNA mixed with 2 µl loading dye on 1 % agarose gel and visualized for the presence of truly band. Samples with high band intensity were used for PCR. The genomic DNA samples with missed and highly smeared bands were re-extracted. Nano-drop 2000 spectrophotometer machine was used to estimate the quality and quantity of the samples using an absorbance of 1 µl at 260 nm, 280 nm and 230 nm. Samples with high quality (A260/A280 ~ 1.8) and concentration were further diluted to 100 ng/µl for PCR.

4.2.3. Primer Selection

SSR primers that showed association with our interest of quality traits in durum wheat genotypes on previously reported studies (Prasad *et al.* (2003), Zhang *et al.* (2008), Ramya *et al.* (2010), Maccaferri *et al.* (2011), Blanco *et al.* (2012), Moradi *et al.* (2014) and Amallah *et al*. (2016)) were selected (Table 2).

Table 2. List of microsatellite markers used in this study.

	Chr	Locus	T^o a		Sequence (5'-3')	Associated trait	Reference
1	1AS	Xwmc24	52	Fw	GTGAGCAATTTTGATTATACTG	TGW	Zhang *et al*. (2008)
				Rv	TACCCTGATGCTGTAATATGTG	TW, GY	Maccaferri *et al.*(2011)
2	1BL	Xbarc240	60	Fw	AGAGGACGCTGAGAACTTTAGAGAA	TW, YP	Zhang *et al.* (2008)
				Rv	GCGATCTTTGTAATGCATGGTGAAC		
3	2AL	Xgwm294	61	Fw	GGATTGGAGTTAAGAGAGAACCG	DH,PLH, TGW	Maccaferri *et al.*(2011)
				Rv	GCAGAGTGATCAATGCCAGA		
4	2BL	Xgwm47.1	61	Fw	TTGCTACCATGCATGACCAT	PLH, TGW	Maccaferri *et al.*(2011)
				Rv	TTCACCTCGATTGAGGTCCT		
5	3AS	Xbarc12	56	Fw	CGACAGAGTGATCACCCAAATATAA	TGW	Maccaferri *et al.*(2011)
				Rv	CATCGGTCTAATTGTCAATGTA		
6	3BS	Xgwm493	60	Fw	TTCCCATAACTAAAACCGCG	DH,PLH, TW, GY	Maccaferri *et al.*(2011)
				Rv	GGAACATCATTTCTGGACTTTG	GPC	Blanco *et al.* (2012)
7	4AL	Xbarc155	59	Fw	GCGAGTATTGACGTCTTATTTTTGAA	PLH, TW	Maccaferri *et al.*(2011)
				Rv	GCGTCATGAATTCTAACAATGTGCATA		
8	4BS	Xwmc617	59	Fw	CCACTAGGAAGAAGGGGAAACT	TGW, Fe content	Ramya *et al.* (2010) Moradi *et al.* (2014)
				Rv	ATCTGGATTACTGGCCAACTGT	GNS	Blanco *et al.* (2012)
9	4BL	Xgwm513	58	Fw	ATCCGTAGCACCTACTGGTCA	GPC	Prasad *et al.*(2003)
				Rv	GGTCTGTTCATGCCACATTG		
10	5AS	Xgwm120	58	Fw	GATCCACCTTCCTCTCTCTC	GPC, GL	Zhang *et al.* (2008)
				Rv	GATTATACTGGTGCCGAAAC		
11	5BL	Xgwm371	63	Fw	GACCAAGATATTCAAACTGGCC	TGW, TW	Ramya *et al.* (2010) Amallah *et al.* (2016)
				Rv	AGCTCAGCTTGCTTGGTACC		
12	6AL	Xwmc256	61	Fw	CCAAATCTTCGAACAAGAACCC	DH, GY	Maccaferri *et al.*(2011)
				Rv	ACCGATCGATGGTGTATACTGA		
13	6BL	Xbarc178	59	Fw	GCGTATTAGCAAAACAGAAGTGAG	TGW	Ramya *et al.* (2010)
				Rv	GCGACTAGTACGAACACCACAAAA		
14	7BS	Xgwm46	60	Fw	GCACGTGAATGGATTGGAC	GPC	Amallah *et al.* (2016)
				Rv	TGACCCAATAGTGGTGGTCA	GY	Maccaferri *et al.*(2011)

Chr is chromosome; T^0a is annealing temperature; S is short chromosomal arm; L is long chromosomal arm; Fw is forward; Rv is reverse; TGW is thousand grains weight; TW is test weight; GY is grain yield; YP is yellow pigment content; DH is days to heading; PLH is plant height; GPC is grain protein content; GNS is grain number per spike and GL is gluten content.

4.2.4. Procedures of Polymerase Chain Reaction (PCR)

Individual PCR amplifications of each microsatellite was carried out in a PCR tube with 10 µl reaction volume containing 7.3 µl PCR water, 1 µl PCR buffer, 0.25 µl dNTPs, 0.15 µl of each forward and reverse primer, 0.15 µl Taq-polymerase enzyme and 1 µl of extracted DNA template of each sample. The amplification of each primer was 4 min preheating and initial denaturing at 94 ^{o}C; 35 cycles of 30 s denaturation at 94 ^{o}C, 30s for primer annealing at a temperature depending on the primer used and 90 s for extension at 72 ^{o}C; and 8 min for final extension at 72 ^{o}C.

4.2.5. Polyacrylamide Gel Electrophoresis (PAGE)

The PCR products were mixed with DNA loading dye and loaded into the wells of 8 % polyacrylamide gel soaked in 0.5x TBE running buffer hold by vertical gel tank. A standard 100 base pair (bp) DNA ladder was used to estimate the molecular weight of amplified products. Electrophoresis was conducted for 3:30 hr. at a constant voltage (110 V) and current (30 mA).

4.2.6. Data Documentation and Scoring

The electrophoresed gels were visualized using Bio rad Gel-Doc EZ gel documentation machine (Bio-Rad Laboratories, Inc.). Microsatellite loci and alleles appear as bands on gels. Fragments with clearly distinguishable and reproducible bands were considered. PCR products with faint bands and negative results were repeated for further confirmation. The molecular weight of fragments in base-pair unit was estimated by Image-Lab 6.1 software (Bio-Rad Laboratories, Inc.) and also scored as presence (1) and absence (0). Each fragment was scored independently based on the expected fragment size of each primer.

4.2.7. Data Analysis

The data generated from data scoring were subjected to further analysis using NTSYS-pc version 2.1 software (Rohlf, 2002), Arliquin version 3.5.2 software (Excoffier and Lischer, 2010), Power-marker version 3.25 software (Liu and Muse, 2005), GenAIEx version 6.5 software (Peakall and Smouse, 2012), MEGA version X software (Kumar *et al.,* 2018) and STRUCTURE version 2.3.4 software (Pritchard *et al*., 2010).

4.2.7.1. Genetic Diversity Indices

Number of alleles per locus (Na), major allelic frequency (MAF), gene diversity (h) and polymorphic information content (PIC) of SSR markers were computed using Power marker version 3.25 software (Liu and Muse, 2005). Mean diversity indices of the 13 populations over 14 SSR loci were computed using GenAlEx version 6.5 software (Peakall and Smouse, 2012).

4.2.7.2. Analysis of Molecular Variance (AMOVA) and F-statistics

AMOVA, that used to partition the total variance among and within populations, and $F_{ST,}$ the proportion of the total genetic variance contained in a sub-population relative to the total genetic variance to determine population differentiation, was computed using Arliquin version 3.5.2 software (Excoffier and Lischer, 2010).

4.2.7.3. Genetic Distance, Cluster and Principal Coordinate Analysis

Allelic frequency of each genotype for each marker was calculated from the 0 and 1 formatted data. Then the allelic frequency data were used to compute genetic distance using NTSYS-pc version 2.1 software (Rohlf, 2002). Principal coordinate analysis was computed using GenAIEx software (Peakall and Smouse, 2012) from the genetic distance matrix. The cluster analysis was

computed using the genetic distance matrix and neighbor joining (NJ) based clustering were constructed using MEGA version X software (Kumar *et al.*, 2018).

4.2.7.4. Population Structure

The Genotypic data were analyzed for Bayesian based population structure using STRUCTURE version 2.3.4 software (Pritchard *et al.*, 2010). The precise number of cluster (K value) were estimated by setting a burn-in periods of 100,000 in a single run followed by 200,000 Markov Chain Monte Carlo (MCMC) replications for K equals to 1 up to 10 using 20 iterations for each K. The optimal K-value was determined by the ΔK method of Evanno *et al.* (2005) using STRUCTURE HARVESTER Web v0.6.94 (Earl and vonHoldt, 2012). Packaging and plotting of the population structure inferences of the optimum K were carried out by using the beta version of CLUMPAK software (Kopelman *et al.*, 2015).

5. Results

5.1. Phenotypic Diversity among Ethiopian Durum Wheat Landraces Based on Quantitative Traits

5.1.1. Performance of Quantitative Traits

5.1.1.1. Ranges of Genotypes over Locations

The ranges of quantitative agro-morphological and quality traits scored at Sinana, Chefe Donsa and combined of the two locations are presented in Table 3. At both environments, wide ranges were scored for majority of studied traits except DH, DM and STD. Combined data over location revealed wide ranges for PLH, NET, TGW, GY and GL. The mean value of GPC, GL and GY over combined locations ranged from 11.45% to 16.15%, 17.5% to 35.38%, and 4.3 t/ha to 12.6 t/ha, respectively (Table 3).

Table 3. Minimum, maximum and mean values of 11 quantitative traits of 420 durum wheat genotypes for Sinana, Chefe Donsa and combined over two locations.

Parameter	Sinana			Chefe Donsa			Combined		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
DH	59	68.01	75.5	64	74.85	82	61.75	71.43	78
DM	134	138.2	145.5	131.5	137.4	147	134	137.8	145.2
SPS	16.35	20.89	25.35	13.7	18.15	22.1	16.48	19.52	22.27
SPL (cm)	5.15	9.5	13.5	5.5	7.7	10.6	5.325	8.6	11.5
PLH (cm)	79	117.1	156	72.5	100.4	131	75.75	108.7	138.5
NET	3.3	6.86	14.3	3.2	6.5	12.4	4.65	6.68	10.82
STD (%)	80	89.4	95	80	89.13	95	82.5	89.26	95
TGW (g)	22.9	42.05	57.1	32.96	42.7	56.2	32.01	42.37	55.4
GY (t/ha)	2.89	8.3	16.5	2.6	8.5	16.5	4.3	8.4	12.6
GL (%)	18.75	28.63	42.25	11.85	22.85	35.2	17.5	25.74	35.38
GPC (%)	11.75	14.66	18.2	9.69	12.09	15.9	11.45	13.37	16.15

DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content. Min is the minimum and Max is the maximum.

5.1.1.2. Mean Performance of Genotypes

The mean performances of 420 Ethiopian durum wheat landraces and cultivars at Sinana, Chefe Donsa and combined over the two locations for all genotypes are presented in Table 3. The mean performance of GL was 28.63% at Sinana, 22.85% at Chefe Donsa and 25.74% over the combined locations. The highest values of GL (above 34%) were obtained from genotypes 150 (35.4%) and 364 (34.8%). The mean performance of GPC was 14.66% at Sinana, 12.09% at Chefe Donsa and 13.37% over the combined locations. The highest values of GPC (above 15.5%) were scored by genotypes 208 (16.15%), 190 (15.85%), 386 (15.8%) and 191 (15.63%). The mean performance of GY was 8.3 t/ha at Sinana, 8.5 t/ha at Chefe Donsa and 8.4 t/ha over the combined locations. Genotypes with the highest GY were 393 (12.64 t/ha), 397 (12.53 t/ha), 413 (12.03 t/ha), 390 (11.75 t/ha), 392 (11.6 t/ha), 416 (11.56 t/ha), 415 (11.46 t/ha), 292 (11.33 t/ha), 420 (11.16 t/ha), 250 (11.1 t/ha) and 293 (11.02 t/ha) and performed well in all locations compared to the other genotypes.

5.1.2. Analysis of Variance

Analysis of variance (ANOVA) with mean square of genotype, replication and block for 11 measured traits at Sinana and Chefe Donsa, and combined locations by including mean square of location are presented in Table 4 and Table 5, respectively. ANOVA revealed highly significant effects on the majority of the quantitative traits with R^2 value ranging from 0.64 to 0.89 over combined locations (Table 5), 0.57 to 0.9 at Sinana and 0.71 to 0.88 at Chefe Donsa (Table 4).

ANOVA revealed highly significant ($P<0.001$) differences for all quantitative traits except DM and STD at Sinana and for all quantitative traits at Chefe Donsa (Table 4) and combined location (Table 5). Test locations have had also a pronounced effects ranging from highly significant

(P<0.001) on nine of the eleven traits and significant (P<0.05) on GY to non-significant on STD (Table 5). Similarly, genotype by location interaction showed highly significant (P<0.001) variation on 7 of the traits and significant (P<0.05) on DH, PLH and DM, except STD (Table 5).

Table 4. Analysis of variance of 11 quantitative traits among 420 Ethiopian durum wheat genotypes tested at Sinana (upper) and Chefe Donsa (bottom) during 2019/20.

Traits		GEN (419)	REP (1)	REP:BLK (40)	Error (379)	Mean	CV (%)	LSD	R^2
DH	SN	20.97***	56.58 **	7.8ns	6.91	68.01	3.83	4.22	0.78
	CDR	25***	385***	15.5***	4.6	74.85	2.88	3.85	0.87
DM	SN	7.442ns	14.933ns	8.9ns	7.1	138.2	1.89	5.29	0.57
	CDR	10***	119***	5.17***	2.6	137.4	1.16	2.71	0.82
SPS	SN	3.230***	14.300**	4.2***	1.82	20.89	6.52	1.66	0.67
	CDR	3.1***	16.8**	7.5***	1.54	18.15	6.91	1.62	0.74
SPL	SN	5.411***	27.97***	6.1***	1.35	9.5	12.5	2	0.83
	CDR	2.0***	177***	2.41***	0.62	7.7	10.4	1.29	0.83
PLH	SN	158.2***	154.29ns	149.8*	99	117.1	8.39	9.41	0.66
	CDR	135***	5306***	121***	33	100.4	5.79	9.99	0.84
NET	SN	3.41***	12.17**	5.38***	1.78	6.86	20	1.5	0.7
	CDR	2.9***	125***	9.6***	1	6.4	15.7	1.37	0.82
STD	SN	12.8ns	340.7***	26.7***	11.1	89.4	3.77	6.75	0.62
	CDR	21***	14.4ns	48.6**	11.4	89.1	3.82	4.09	0.71
GY	SN	737.8***	10223***	484***	239	83.4	18.5	25	0.79
	CDR	442***	9735***	443***	154	85	14.6	19.2	0.79
TGW	SN	49.34***	17.46ns	15.3***	5.97	42.1	5.9	4.7	0.9
	CDR	44***	559***	21***	6.9	42.7	6.17	4.85	0.88
GL	SN	26.6***	427.7***	66.4***	6.1	28.63	8.76	4.33	0.86
	CDR	23***	82.08**	25.4***	8.88	22.9	13.1	4.41	0.76
GPC	SN	2.84***	55.57***	8.38***	0.77	14.7	6.1	1.4	0.8
	CDR	1.7***	1.42ns	1.42***	0.73	12.1	7.07	1.21	0.74

DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content. GEN is genotype; REP is replication; BLK is block; CV is coefficient of variation; LSD is least significant difference; and R^2 is regression coefficient. SN is Sinana and CDR is Chefe Donsa. ***, ** and * are significant at p<0.001, p<0.01 and p<0.05, respectively. ns is non-significant at p<0.05.

Table 5. Combined analysis of variance of 11 phenotypic quantitative traits among 420 Ethiopian wheat genotypes tested at Sinana and Chefe Donsa during 2019/20.

Trait	Model (880)	LOC (1)	REP (1)	BLK (40)	GEN (419)	GEN:LOC (419)	Error	CV	LSD	R^2
DH	44.9 ***	19632***	73***	89***	31***	8*	6.0	3.3	3.36	0.89
DM	9.1***	298.37***	24.29*	26.9***	9.6***	6.14*	5.1	1.2	2.71	0.67
SPS	6.96***	3159.8***	31.1***	9.4***	3.7***	2.4***	1.8	6.7	1.35	0.81
SPL	5.43***	1358.8***	173***	11.9***	4.9***	1.7***	1.2	11.7	1.37	0.83
PLH	283***	116913***	3635***	612***	165***	83*	72.0	7.4	9.33	0.81
NET	3.62***	54.14***	107***	18.9***	2.8***	2.7***	1.7	18.0	1.11	0.71
STD	18.8***	27.51ns	248***	92.3***	17***	13.2ns	11.8	3.8	4.09	0.64
TGW	45.8***	172.7***	387***	138***	68***	14***	7.4	6.1	4.67	0.87
GY	586***	1082.4*	3ns	1809***	639***	416***	233.0	16.7	18.10	0.74
GL	41.6***	14050***	68*	85***	30***	16***	10.0	10.8	3.82	0.82
GPC	5.6***	2774.2***	19.6***	12.3***	2.3***	1.6***	1.0	6.6	1.09	0.86

DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content. GEN is genotype, LOC is location; REP is replication; BLK is block; CV is coefficient of variation; LSD is least significant difference; and R^2 is regression coefficient. ***, ** and * are significant at $p<0.001$, $p<0.01$ and $p<0.05$, respectively. ns is non-significant at $p<0.05$.

5.1.3. Estimation of Components of Variation

5.1.3.1. Genotypic and Phenotypic Coefficient of Variation

The combined analysis of genotypic variance, environmental variance, genotype by location interaction variance, phenotypic variance, genotypic and phenotypic coefficient of variation, heritability, genetic advance and genetic advance as percent of mean of 11 quantitative traits considering combined locations are presented in Table 6. Both medium GCV and PCV values were observed for SPL (11.2 % and 13.5 %, respectively) and GY (12 % and 17.1 %, respectively), while the remaining traits showed low values of GCV and PCV except the value of PCV for NET (14.5 %), TGW (10.2 %) and GL (11.6 %) (Table 6).

5.1.3.2. Broad Sense Heritability and Genetic Advance

Heritability in broad sense estimated for 11 quantitative traits of 420 durum wheat genotypes based on combined locations is indicated in Table 6. The value of heritability was high (>50%), intermediate (40-50%) and low (<40%) for the all traits it ranged from 27.7 % for STD to 81 % for TGW. Within this range, the heritability values were high for TGW (81 %), DH (76.5 %), SPL (68.5 %), GL (55.6) and PLH (52.9 %), intermediate for GY (49.4 %), GPC (44.8 %), SPS (44.2 %), DM (42.6 %), NET (29.5 %) and STD (27.7 %). In our study, the GAM estimated for 11 quantitative traits considering combined location ranged from 1.1 % (DM) to 19.1 % (SPL). Among all quantitative traits evaluated, higher GAM recorded for SPL (19.1), GY (17.4), TGW (17.0) and GL (13.4) and the remaining quantitative traits showed low GAM value.

Table 6. Mean, genotypic variance (σ_g^2), environmental variance (σ_e^2), phenotypic variance (σ_p^2), genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), heritability in broad-sense (H_b^2), genetic advance (GA) and genetic advance as percent of mean (GAM) of studied durum wheat genotypes for combined locations.

Traits	Mean	σ_e^2	σ_{g*l}^2	σ_g^2	σ_p^2	GCV	PCV	H_b^2(%)	GA	GAM
DH	71.4	6.0	0.79	6.16	8.1	3.5	3.97	76.5	4.47	6.3
DM	137.8	5.1	0.55	1.14	2.7	0.8	1.19	42.6	1.44	1.1
SPS	19.5	1.8	0.30	0.48	1.1	3.5	5.31	44.2	0.94	4.8
SPL	8.6	1.2	0.25	0.93	1.4	11.2	13.50	68.5	1.64	19.1
PLH	108.7	72.0	5.50	23.25	44.0	4.5	6.10	52.9	7.22	6.7
NET	6.7	1.7	0.50	0.28	0.9	7.9	14.50	29.5	0.59	8.8
STD	89.3	11.8	0.70	1.26	4.6	1.3	2.40	27.7	1.22	1.4
TGW	42.4	7.4	3.35	15.03	18.6	9.2	10.20	81.0	7.19	17.0
GY	84.1	233.0	91.50	101.60	205.6	12.0	17.10	49.4	14.59	17.4
GL	25.7	10.0	3.00	5.00	9.0	8.7	11.60	55.6	3.43	13.4
GPC	13.4	1.0	0.30	0.33	0.7	4.3	6.40	44.8	0.79	5.9

DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content.

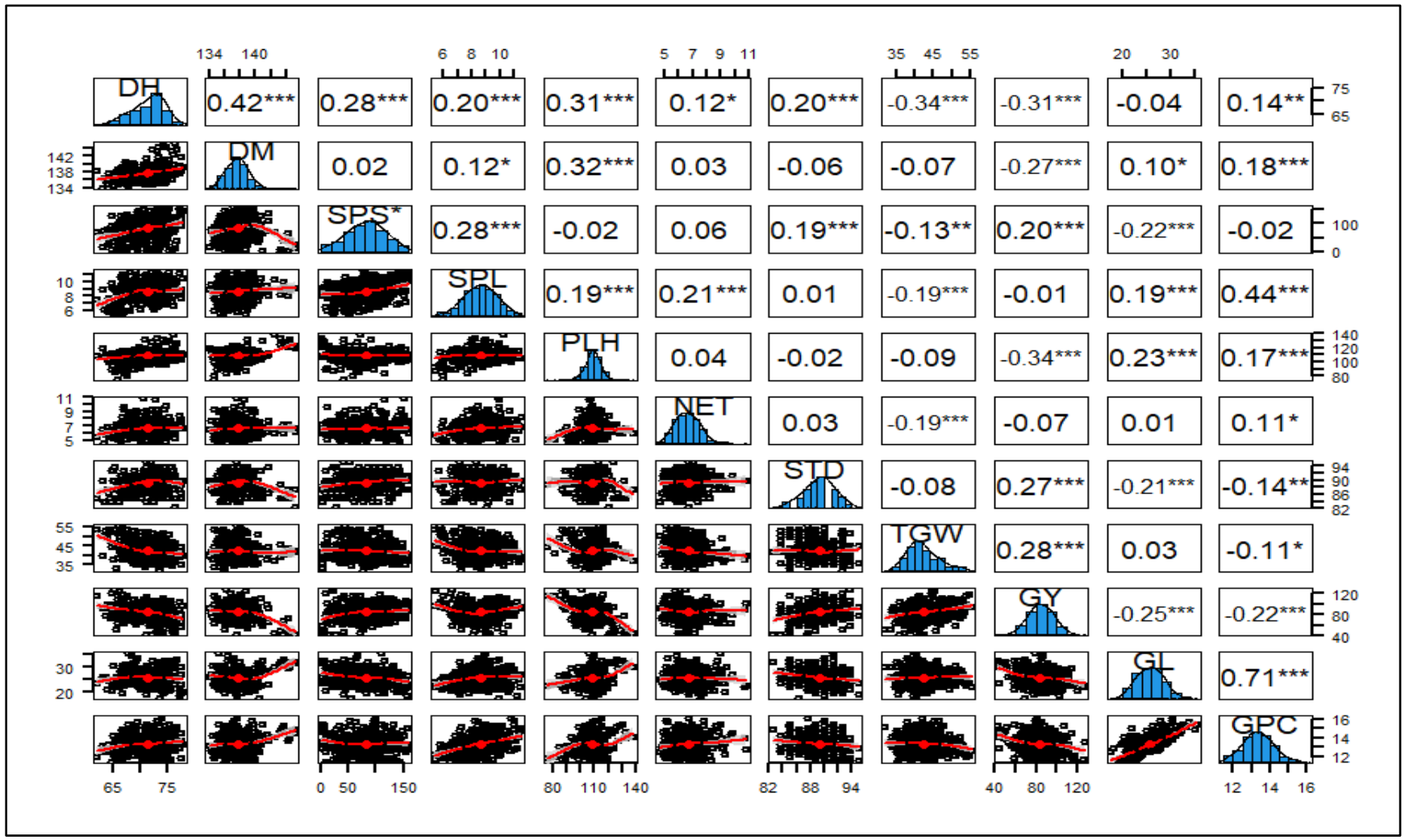
5.1.4. Correlation Analysis among Phenotypic Traits

The phenotypic correlation coefficients of all possible pairs of quantitative traits showed significant association among most traits (Table 7) and described in the following subsections.

5.1.4.1. Protein Related Traits with Other Traits

Gluten content revealed positive phenotypic correlation with GPC (0.71***) and SPL (0.19***), PLH (0.23***) and DM (0.10*). On the other hand, GL showed negative phenotypic correlation with SPS (-0.22***), STD (-0.21***) and GY (-0.25***). Positive correlation of GPC was observed with DM (0.18***), SPL (0.44***), PLH (0.17***), DH (0.14**) and NET (0.11*). In contrast, GPC showed negative correlation with GY (-0.22***), STD (-0.14**) and TGW (-0.11*) (Table 7).

Table 7. Phenotypic correlation coefficients (above diagonal), distribution (diagonal) and scatterplot with regression curve of 11 quantitative traits among 420 Ethiopian durum wheat genotypes.



DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content.

5.1.4.2. Yield Traits with Other Traits

The GY showed positive phenotypic correlation coefficient with TGW (0.28***), SPS (0.20***) and STD (0.27***), negative correlation with DH (-0.31***), DM (-0.27***), PLH (-0.34***), GL (-0.25***), GPC (-0.22***), and non-significant correlation with SPL (-0.01ns) and NET (-0.07ns). TGW noted negative phenotypic correlation for traits DH (-0.34***), SPS (-0.13***), SPL (-0.19***), NET (-0.19***) and GPC (-0.11*). The only positive phenotypic correlation was found with GY (0.28***). SPS signified positive phenotypic correlation coefficient with DH (0.28***), SPL (0.28***), STD (0.19***) and GY (0.2***); negative correlation with TGW (-0.13**) and GL (-0.22**); and non-significant correlation with remaining traits (Table 7).

5.1.4.3. Agronomic Traits with Other Traits

Spike length showed positive phenotypic correlation coefficient with DH (0.2***), SPS (0.28***), PLH (0.19***), NET (0.21***), GL (0.19***), GPC (0.44***) and DM (0.12*). Conversely, the only negative correlation of SPL was recorded with TGW (-0.19***). PLH signified positive phenotypic correlation coefficient with DH (0.31***), DM (0.32***), SPL (0.19***), GL (0.23***) and GPC (0.17***); negative correlation with GY (-0.34); and non-significant correlation with the remaining traits (Table 7).

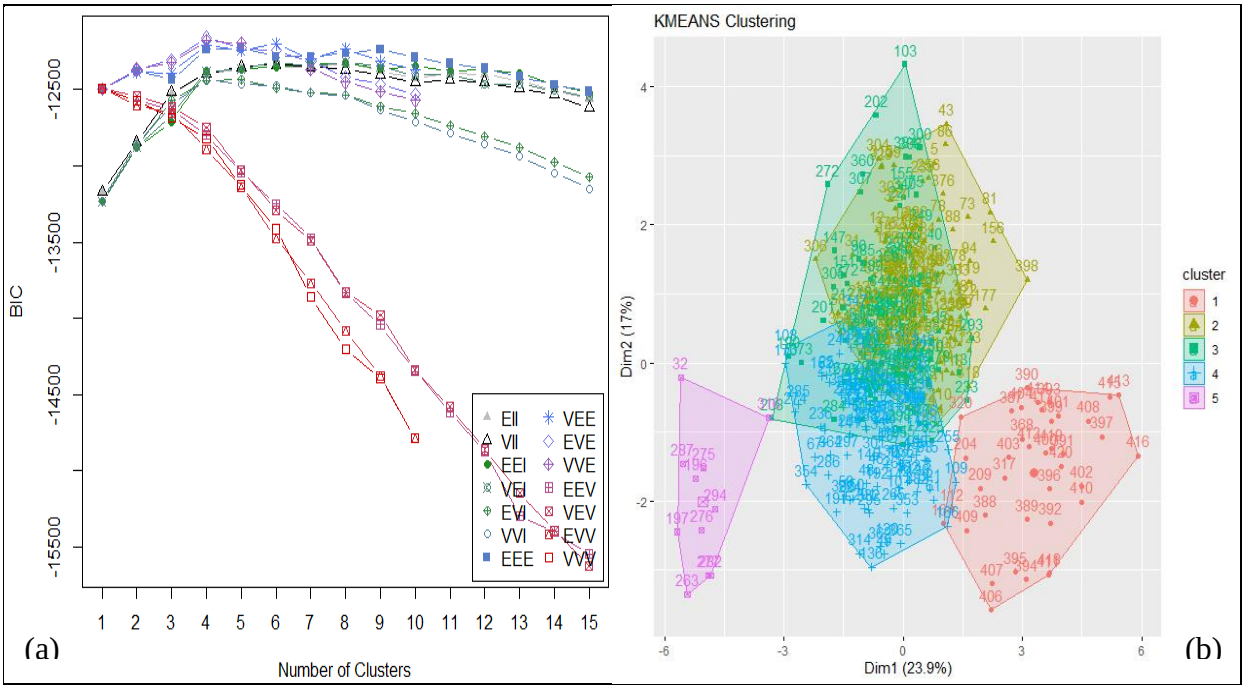
5.1.4.4. Phenological Traits with Other Traits

Days to heading revealed positive phenotypic correlation with DM (0.42***), SPS (0.28***), SPL (0.20***), PLH (0.31***), STD (0.20***), GPC (0.14**) and NET (0.12*); negative correlation with TGW (-0.34***), GY (-0.31***); and non-significant correlation with GL (-0.04ns). DM showed positive phenotypic correlation with DH (0.42***), PLH (0.32***), GPC (0.18***), SPL (0.12*) and GL (0.10*); negative correlation with GY (-0.31***); and non-significant correlation with the remaining traits (Table 7).

5.1.5. Cluster Analysis

5.1.5.1. Clustering of Genotypes

The cluster analysis was computed to assess the genetic relationship among 420 Ethiopian durum wheat genotypes using quantitative phenotypic traits (Table 8 and Figure 3). The estimation of optimal number of clusters based on Bayesian Information Criterion (BIC) method identified five clusters as more appropriate number of clusters (Figure 2).



Dim1 and Dim2 are dimension 1 and 2, respectively. EII is Spherical Equal Equal, VII is Spherical Variable Equal, EEI is Diagonal Equal Equal Coordinate axes, VEI is Diagonal Variable Equal Coordinate axes, EVI is Diagonal Equal Variable Coordinate axes, VVI is Diagonal Variable Variable Coordinate axes, EEE is Ellipsoidal Equal Equal Equal, EVE is Ellipsoidal Equal Variable Equal, VEE is Ellipsoidal Variable Equal Equal, VVE is Ellipsoidal Variable Variable Equal, EEV is Ellipsoidal Equal Equal Variable, VEV is Ellipsoidal Variable Equal Variable, EVV is Ellipsoidal Equal Variable Variable, VVV is Ellipsoidal Variable Variable Variable

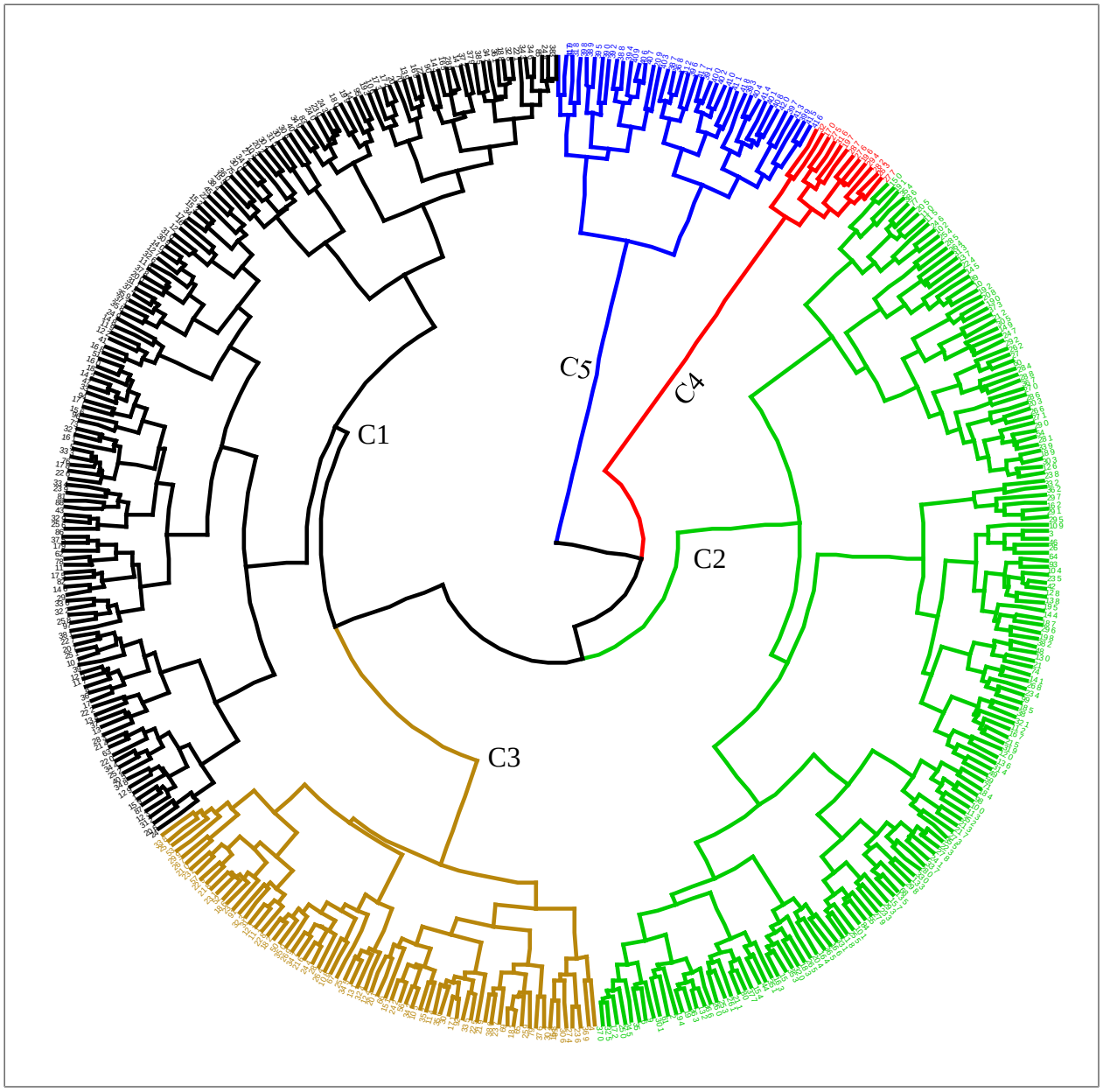
Figure 2. Biplot of the number of clusters with corresponding BIC value (a) and K-means clustering of studied genotypes in to five groups on the two-dimensional plane (b).

Tested genotypes were grouped into five main clusters: C1, C2, C3, C4 and C5 at a genetic distance of 18.6 with significant chi-square values in all inter-cluster distances. The number of genotypes per cluster ranged from 12 (C4) to 155 (C2) (Table 8).

Table 8. The number of genotypes accompanied by each cluster

Cluster	Number of genotypes	Code of genotypes
1	148	1, 5, 7, 8, 11, 12, 13, 14, 16, 17, 18, 20, 22, 24, 29, 33, 34, 36, 37, 39, 40, 41, 43, 45, 47, 51, 62, 63, 70, 72, 73, 75, 76, 78, 81, 82, 83, 85, 86, 87, 88, 90, 94, 95, 96, 97, 103, 106, 107, 116, 117, 119, 122, 123, 125, 127, 129, 134, 139, 140, 142, 143, 146, 147, 148, 152, 153, 155, 156, 159, 160, 161, 165, 168, 169, 173, 174, 175, 176, 177, 178, 179, 180, 188, 193, 199, 200, 202, 206, 207, 214, 220, 221, 222, 226, 227, 230, 239, 240, 243, 248, 249, 251, 252, 256, 258, 272, 283, 289, 300, 304, 307, 308, 309, 310, 311, 312, 315, 322, 326, 327, 328, 329, 333, 334, 336, 340, 341, 343, 344, 346, 348, 349, 358, 359, 360, 361, 371, 372, 374, 375, 378, 379, 381, 383, 384, 385, 405
2	155	2, 3, 6, 9, 10, 15, 19, 21, 26, 27, 35, 38, 42, 44, 46, 48, 49, 54, 55, 59, 60, 61, 64, 68, 71, 74, 77, 80, 84, 89, 93, 98, 99, 100, 101, 102, 104, 109, 110, 112, 113, 115, 118, 120, 126, 128, 130, 132, 133, 135, 136, 137, 138, 141, 144, 145, 149, 154, 157, 162, 163, 164, 166, 167, 170, 172, 183, 184, 185, 187, 189, 190, 191, 192, 194, 195, 198, 203, 204, 205, 208, 211, 212, 213, 215, 217, 223, 224, 233, 234, 235, 238, 241, 245, 246, 247, 250, 253, 254, 255, 260, 261, 262, 265, 266, 268, 270, 271, 273, 278, 279, 280, 281, 284, 285, 286, 288, 290, 291, 292, 293, 295, 296, 297, 298, 301, 303, 305, 313, 314, 319, 320, 321, 325, 330, 331, 332, 337, 339, 351, 352, 353, 354, 355, 362, 363, 364, 365, 366, 367, 370, 373, 377, 382, 386
3	68	4, 23, 25, 28, 30, 31, 50, 52, 53, 56, 57, 58, 65, 66, 67, 69, 79, 91, 92, 105, 108, 111, 114, 121, 124, 131, 151, 158, 171, 181, 182, 186, 201, 210, 216, 218, 219, 225, 228, 229, 231, 232, 236, 237, 242, 244, 257, 259, 264, 267, 269, 274, 299, 302, 306, 323, 324, 335, 338, 342, 345, 347, 350, 356, 357, 369, 376, 380
4	12	32, 150, 196, 197, 263, 275, 276, 277, 282, 287, 294, 316
5	37	209, 317, 318, 368, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420

The boxes of this table are colored based on the color of the dendrogram (Figure 3).



C1, C2, C3, C4 and C5 are cluster one, two, three, four and five.

Figure 3. Dendrogram shows the clustering patterns of the 420 Ethiopian durum wheat genotypes.

5.1.5.2. Genetic Distance between Clusters

The magnitude of inter-cluster and intra-cluster distance identifies the extent of genetic diversity among genotypes of different clusters and same cluster, respectively (Table 9). Genotypes found in the same cluster are characterized by very low genetic distance and vise-versa. The intra-

cluster distance was zero for all compared clusters. However, the inter-cluster distance values ranged from 7.1 (between C1 and C2) to 58.7 (between C4 and C5) (Table 9). The lower or the higher values of inter-cluster distance suggest the closer or the wider, respectively, relationship among individuals of a cluster to the other clusters. The second and the third most divergent clusters were C1 and C4 (D^2 = 41.7) and C2 and C4 (D^2 = 38.4), respectively (Table 9).

Table 9. Average inter-cluster distance value of five clusters

Cluster	C1	C2	C3	C4	C5
C1	0				
C2	7.1b	0			
C3	16.7a	13.5a	0		
C4	41.7a	38.4a	30a	0	
C5	20.9a	21.3a	32a	58.7a	0

C1, C2, C3, C4 and C5 are cluster one, two, three, four and five. The boxes of this table are colored based on the color of the cluster on the dendrogram (Figure 3). a = highly significant (P<0.001), b = significant (P<0.01)

5.1.5.3. Mean Performance of Genotypes in Clusters

The mean values of the quantitative phenotypic traits for each cluster showed the existence of significant variation among clusters for individual or multiple traits considered (Table 10).

Table 10. The mean value of all quantitative traits in each cluster.

Cluster	DH	DM	SPS	SPL	PLH	NET	STD	TGW	GY	GL	GPC
C1	72.7	137.9	19.9	8.73	110.3	6.7	90.5	40.7	88.1	24	13
C2	70.3	137.4	19.3	9.03	109.1	6.72	88.6	43.7	84.2	27.3	13.8
C3	73.6	138.2	19.7	8.25	107.1	6.7	89.2	39.7	72.2	26.2	13.6
C4	75.2	143.3	17.9	9.14	127.5	6.5	85.8	41.2	52	31.1	14.6
C5	66.1	136.5	19	6.72	98	5.9	88.3	49.1	100	23.8	12.5

DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content. C1, C2, C3, C4 and C5 are cluster one, two, three, four and five. The cluster boxes in this table are colored based on the color of the dendrogram (Figure 3).

The C1 comprised 148 genotypes i.e. 35% of the genotypes studied. The mean performance of genotypes contained in C1 was lower for GL, TGW and GPC, whilst higher for STD, SPS and GY (Table 10). Best five high-performing genotype of this cluster were: 308, 103, 307, 73 and 272 (Table 11). Cluster 2 is the major cluster which consists of 155 (37% of studied genotypes) genotypes. C2 is characterized by having highest value for NET, higher value for GL, GPC, TGW and SPL, and lower value for DH and DM (Table 10). Best five high-performing genotypes of this cluster were 364, 297, 362, 145 and 109 (Table 11).

Table 11. The name of five superiorly and inferiorly performed genotypes in each cluster.

Performance level	Cluster				
	C1	C2	C3	C4	C5
Superior genotypes	308	364	25	263	416
	103	297	267	197	397
	307	362	306	277	415
	73	145	299	150	413
	272	109	210	276	402
Inferior genotypes	127	135	171	275	400
	326	157	28	196	414
	348	235	218	32	391
	359	339	182	294	404
	346	19	121	276	412

C1, C2, C3, C4 and C5 are cluster one, two, three, four and five. The boxes of this table are colored based on the color of the cluster on the dendrogram (Figure 3).

Cluster 3, consisting of 68 genotypes, is characterized by lowest value of TGW and lower values of PLH, SPL and GY, while higher values were revealed for DH, DM, STD and SPS (Table 10). Best five high-performing genotypes of this cluster were 25, 267, 306, 299 and 210 (Table 11). Cluster 4, consists of 12 genotypes, characterized by having the highest mean value for GL, GPC, DH, DM, SPL and PLH, and the lowest mean value for GY, SPS and STD (Table 10). Best five high-performing genotypes of this cluster were 263, 197, 277, 150 and 276 (Table 11). Cluster 5, consisting of 37 genotypes, were constituted from cultivars (except 386 and 405) and

landraces 209, 317, 318 and 368. C5 is characterized by having the lowest value for GL, GPC, DH, DM, SPL, PLH and NET, and highest value for TGW and GY (Table 10). Best five high-performing genotypes of this cluster were 416, 397, 415, 413 and 402 (Table 11).

5.1.6. Principal Component Analysis

In the present investigation, the principal component analysis (PCA) was computed to determine the relative contribution of traits accounting for the total variation of studied wheat genotypes. The first four principal components (PCs) with Eigenvalues greater than 1 accounted for 64% of the entire variability (Table 12).

Table 12. Eigenvalues, percent proportion of variance, cumulative percent and eigenvectors of the first four principal components of 11 quantitative phenotypic traits of studied wheat genotypes.

Parameters	PC1	PC2	PC3	PC4
Eigen values	2.63	1.87	1.42	1.1
Proportion of Variance (%)	24	17	13	9.8
Cumulative Proportion (%)	24	41	54	63.8
Traits	Eigen vectors			
Days to heading (DH)	-0.34	0.42	-0.23	0.12
Days to maturity (DM)	-0.33	0.1	-0.32	0.36
Number of spikelet per spike (SPS)	-0.01	0.5	0.28	0.18
Spike length (SPL)	-0.32	0.18	0.50	0.06
Plant height (PLH)	-0.36	0.02	-0.24	0.26
Number of effective tiller (NET)	-0.16	0.18	0.21	-0.57
Stand growth (STD)	0.10	0.42	0.1	0.22
Thousand grains weight (TGW)	0.24	-0.3	0.08	0.53
Grain yield (GY)	0.38	0.14	0.42	0.28
Gluten content (GL)	-0.35	-0.43	0.27	0.1
Grain protein content (GPC)	-0.43	-0.23	0.41	0.08

PC1, PC2, PC3 and PC4 is principal component one, two, three and four, respectively.

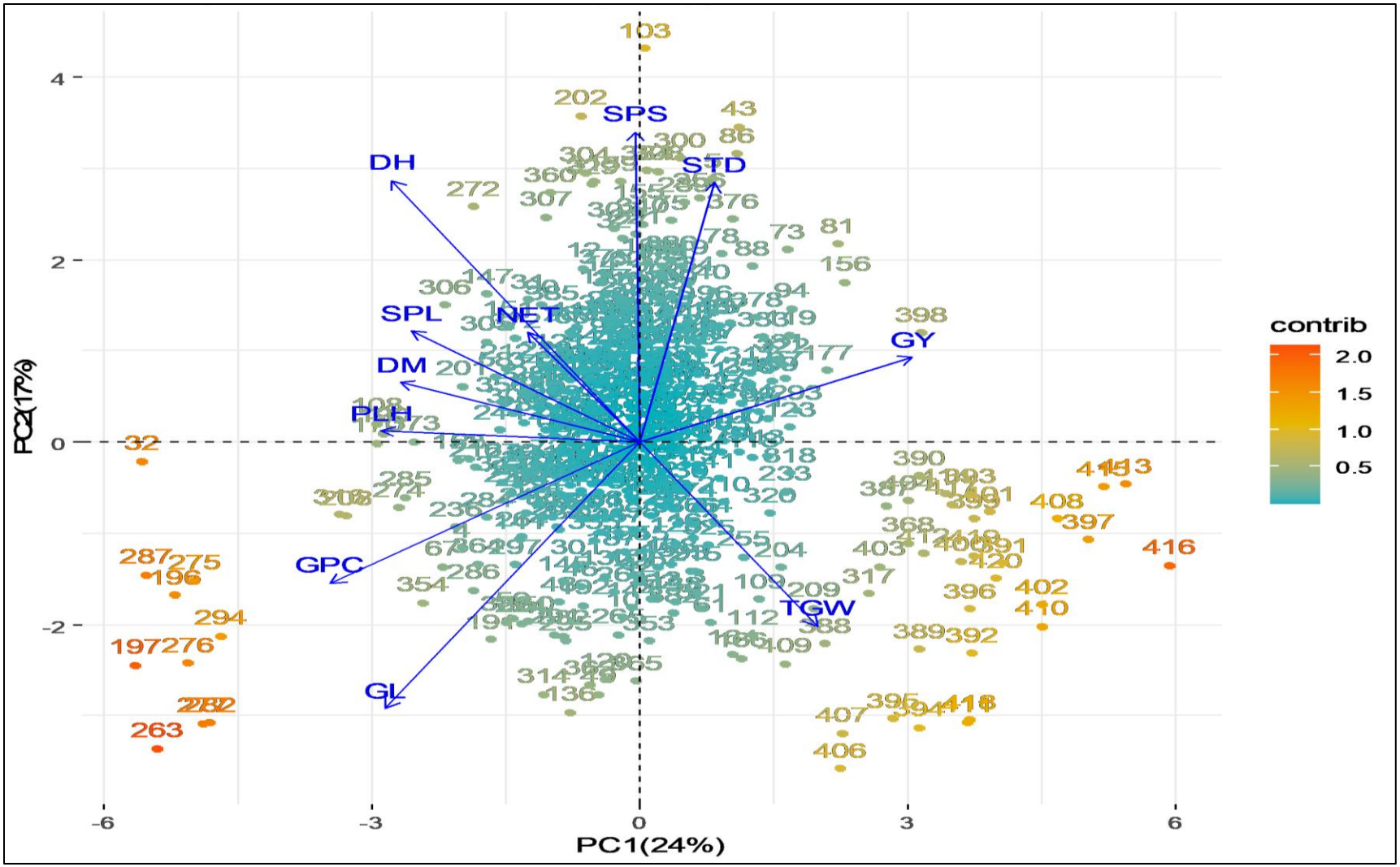
Phenotypic traits: DH, DM, SPL, PLH, GY, GL and GPC contributed much for the principal component 1 (PC1) which explained 24% of the total variation (Table 12). In the principal component 2 (PC2), which explained 17% of the total variation, DH, SPS, STD, TGW and GL had large contribution. DM, SPL, GY and GPC contributed much to the variation in principle component 3 (PC3) which constituted 13% of the total variation. NET and TGW contributed most to principal component 4 (PC4) which signified for 9.8% of the total variation (Table 12).

The bar plot of the first two principal components with their relative contribution estimated from the 11 quantitative traits is presented in Figure 4. The contribution of PCs declines linearly or gradually from PC1 (24%) to PC2 (17%) and went down until PC11 (2%), which suggests the first few principal components had the greatest contribution to the total variation in the Ethiopian durum wheat genotypes for the 11 quantitative traits.

Table 13. Distance and Eigen values of the first four principal components of geographical origins.

Geographical origins	Distance	PC1	PC2	PC3	PC4
Arsi	1.4	1.14	-0.53	0.42	-0.18
Bale-Hararge	0.66	0.18	-0.23	0.25	0.12
East Shewa	0.49	0.4	0.24	0.07	0.1
Gojam	0.9	0.29	-0.01	-0.42	-0.46
Gondar-Wello	0.5	-0.09	0.11	-0.24	-0.24
North Shewa	0.63	0.1	0.54	0.06	0.17
SNNP	0.6	0.1	0.06	-0.23	-0.33
Tigray	0.69	0.15	0.11	-0.1	0.02
Cultivars	3.86	-3.49	-1.46	-0.16	0.35
West Shewa	0.89	0.22	0.77	-0.12	0.09

PC1, PC2, PC3 and PC4 is principal component one, two, three and four, respectively.



DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content. PC1 and PC2 are principal component one and two, respectively. Contrib is contribution.

Figure 4. Principal component biplot of 420 Ethiopian durum wheat genotypes with quantitative phenotypic traits.

The distribution of landraces along the first two PCs was concentrated near to the center except for the cultivars and some landraces from Arsi and North Shewa (Table 13). The extremes of PC1 axis were occupied by landraces from North Shewa whereas the extreme of PC2 axis is occupied by landraces from Arsi and cultivars with negative principal scores (Table 13). Overlaying traits on genotypes showed that yield traits (mainly TGW and GY) are more important for discriminating cultivars from landraces. In general, DH, SPS, GL, TGW, GPC and GY were the most important traits for discriminating genotypes.

5.2. Phenotypic Diversity among Ethiopian Durum Wheat Landraces Based on Qualitative Traits

Six qualitative traits were recorded from studied genotypes and analyzed to determine the magnitude and pattern of diversity among all genotypes and their geographical origin. The recorded qualitative traits were reorganized to make frequency table of each traits for each genotypes. Then the qualitative frequency data were analyzed for chi-square test, Shannon diversity index and correspondence analysis.

5.2.1. Frequency Distribution and Chi-square Test of Qualitative Traits

The frequency distribution of qualitative traits expressed in percentage and the result of computed chi-square test for each geographical origin of genotypes are presented in Table 14. All of the six qualitative traits with their different phenotypic classes has distributed among geographical origins. The hairless glume phenotypic class was predominant in all geographical origins and cultivars compared to all phenotypic classes of qualitative traits. In addition, the less hairy and black glume color phenotypic classes were the rarest in all geographical origins and cultivars compared to all phenotypic classes of qualitative traits.

Table 14. Frequency distribution and chi-square values (X^2) based on geographical origins.

Phenotypic class	Gojam	Bale-Hararge	Tigray	Gondar-Wello	West Shewa	Arsi	East Shewa	Cultivars	North Shewa	SNNP	Mean
No lodging	13.6	14.3	16.7	15	14.6	21	10.1	58.1	10.4	11.8	18.6
Low lodging	37.1	42.3	29.2	31.9	36.5	39.8	37.8	21.3	35.8	31.6	34.3
Medium Lodging	41.4	38.7	49	47.5	40.6	35.8	45.7	17.6	48.5	51.3	41.6
Complete Lodging	7.9	4.8	5.2	5.6	8.3	3.4	6.4	2.9	5.4	5.3	5.5
X^2	920	888	546	817	566	1240	2200	565	1308	347	
SPD_Lax	30	38.1	29.2	33.8	26	45.5	39.1	0.7	32.3	35.5	31
SPD_Inter-mediate	57.9	54.2	55.2	56.3	40.6	48.3	42	8.1	47.3	52.6	46.3
SPD_Dense	12.1	7.7	15.6	10	33.3	6.3	18.9	91.2	20.4	11.8	22.7
X^2	1337	1493	796	1593	1169	1475	3578	657	2376	568	
GLC_White	72.1	50.6	57.3	63.8	62.5	50	61.4	80.1	64.6	51.3	61.4
GLC_Red-brown	26.4	44.6	40.6	31.9	33.3	49.4	31.6	14.7	30.8	48.7	35.2
GLC_Black	1.4	4.8	2.1	4.4	4.2	0.6	6.9	5.2	4.6	0	3.4
X^2	1307	1402	1037	1442	973	1502	3929	1347	252	492	
Hairless	93.6	97.6	90.6	88.8	94.8	93.2	92.3	95.6	88.5	89.5	92.5
Less hairy	0	1.2	1	2.5	0	0	1.1	1.5	1.5	2.6	1.1
Highly hairy	6.4	1.2	8.3	8.8	5.2	6.8	6.7	2.9	10	7.9	6.4
X^2	924.2	1037	993.1	1367	504.6	1139	3215	673	2365	442	
KCL_White	87.1	62.5	81.3	78.8	68.8	63.1	58.8	100	63.8	72.4	73.7
KCL_Brown	5.7	20.2	4.2	5.6	10.4	15.3	19.4	0	14.2	7.9	10.3
KCL_Purle	7.1	17.3	14.6	15.6	20.8	21.6	21.8	0	22	19.7	16.1
X^2	1138	1818	1018	1646	1164	2061	4203	0	2612	706	
Non-vitreous	10	7.1	6.3	5.6	6.3	1.1	3.2	2.9	7.3	18.4	6.8
Partly vitreous	17.1	20.2	22.9	17.5	12.5	8	19	20.6	15.8	31.6	18.5
Vitreous	72.9	72.6	70.8	76.9	81.3	91	78	76.5	76.9	50	74.7
X^2	1328	1284	732	1197	728	1412	3108	1109	2567	560	

SPD is spike density, GLC is glume color and KCL is kernel color.

The vitreous kernel phenotypic class, a trait that is among the most important end-user quality traits, was the second predominant trait observed in all geographical origins and in cultivars. The highest vitreous kernel frequency recorded from Arsi geographic origin (91 %) and followed by West Shewa geographic origin (81.3 %) (Table 14).

The other important quality trait in the processing sector, white kernel color was the third common phenotypic class in all geographical origins and cultivars. All cultivars exhibited white kernel color. Phenotypic classes in SPD were evenly distributed among geographical regions. However, intermediate and lax spike density was common in Ethiopian landraces, while nearly all cultivars exhibited dense SPD. Low lodging and medium lodging phenotypic classes of LDG were common in Ethiopian landraces and no lodging was common in cultivars as they are short in plant height. Among landraces, complete lodging was lower in Arsi collection. White glume color was dominant in all geographical origins and cultivars. Black glume color was rare in all geographical origins and cultivars, moreover it is absent in the SNNP collection.

The chi-square test revealed highly significant effects of genotypes on most of the qualitative traits with Shannon diversity index value ranging from 0.27-0.96 over combined locations, 0.26-0.96 at Sinana and Chefe Donsa (Table 15). Genotypes showed a highly significant chi-square value for all qualitative traits except VTR at Sinana, Chefe Donsa and combined locations. The test locations had also a pronounced effect ranging from highly significant chi-square value for LDG and GLC to non-significant chi-square value for SPD, GLH, KCL and VTR.

Table 15. The chi-square value (X^2) and Shannon diversity index value (H') of qualitative traits at Sinana, Chefe Donsa and combined over locations.

Traits	Observed phenotypic Class	Class	SN				CDR				Combined			
			Proportion (%)	X^2		H'	Proportion (%)	X^2		H'	Proportion (%)	X^2		H'
				GEN	REP			GEN	REP			GEN	LOC	
LDG	No lodging	4	19.2	1576 ***	14.1 **	0.9	14.9	1327 ns	47.8 ***	0.81	17.0	1638 ***	32.9 ***	0.87
	Low lodging		31.8				38.9				35.4			
	Medium lodging		40.8				43.3				42.0			
	Complete lodging		8.2				2.9				5.6			
SPD	Lax	3	34.8	1229 ***	11.7 **	0.96	31	1246 ***	26.2 ***	0.96	32.9	2059 ***	3.24 ns	0.96
	Inter-mediate		45.1				46.4				45.8			
	Dense		20.1				22.6				21.3			
GLC	White	3	57.6	1255 ***	0.21 ns	0.78	66	1259 ***	0.09 ns	0.65	61.8	1974 ***	26.5 ***	0.73
	Red-brown		36.1				32.1				34.1			
	Black		6.3				1.9				4.1			
GLH	Hairless	3	92.4	1154 ***	2.75 ns	0.26	92.1	1129 ***	1.7 ns	0.26	92.3	1722 ***	5.1 ns	0.27
	Less		0.5				1.8				1.1			
	High		7.1				6.1				6.6			
KCL	White	3	70.1	1128 ***	65 ***	0.74	70.6	1102 ***	66 ***	0.74	70.3	2205 ***	0.05 ns	0.74
	Brown		12.5				12.3				12.4			
	Purple		17.4				17.1				17.3			
VTR	Non-vitreous	3	5.8	901 ns	59.7 ***	0.62	5.8	867 ns	60 ***	0.62	5.8	1744 ***	0.02 ns	0.62
	Partly-vitreous		17.9				17.7				17.8			
	Vitreous		76.3				76.5				76.4			

LDG is lodging; SPD is spike density; GLC is glume color; GLH is glume hairiness, KCL is kernel color, VTR is vitreousness, SN is Sinana; CDR is Chefe Donsa, GEN is genotype; REP is replication; and LOC is location. ***, ** and * are significant at p<0.001, p<0.01 and p<0.05, respectively. ns is non-significant at p<0.05.

5.2.2. Shannon Diversity Index of Qualitative Traits

Among the six qualitative traits, LDG has exhibited an intermediate and the highest mean diversity index value (0.46) across all geographical origins followed by SPD (0.35) and VTR (0.25). The SNNP collection scored the highest H' value for LDG (0.52), SPD (0.50), GLH (0.14) and VTR (0.46) than other geographical origins. The H’ value for KCL was below 0.20 with the highest for North Shewa collection (0.20). The H’ value of GLH with the mean value 0.06 was closer to zero across all geographical origins except the maximum value from SNNP (0.14) (Table 16).

Table 16. Mean Shannon diversity index (H') for each qualitative trait across geographical origin.

Geographical origin	LDGH'	SPDH'	GLCH'	GLHH'	KCLH'	VTRH'	Mean
Arsi	0.48	0.36	0.16	0.04	0.15	0.10	0.22
Bale-Hararge	0.47	0.38	0.24	0.03	0.17	0.28	0.26
East Shewa	0.42	0.39	0.18	0.05	0.18	0.22	0.24
Gojam	0.41	0.34	0.09	0.03	0.11	0.22	0.2
Gondar-Wello	0.44	0.33	0.19	0.09	0.09	0.26	0.23
North Shewa	0.45	0.40	0.18	0.06	0.20	0.2	0.25
SNNP	0.52	0.50	0.15	0.14	0.19	0.46	0.33
Tigray	0.47	0.40	0.14	0.02	0.08	0.32	0.24
Cultivars	0.48	0.14	0.12	0.08	0.00	0.23	0.18
West Shewa	0.44	0.24	0.13	0.05	0.13	0.21	0.2
Mean	0.46	0.35	0.16	0.06	0.13	0.25	**0.24**

LDGH’, SPDH’, GLCH’, GLHH’, KCLH’, VTRH’ is the Shannon diversity index of lodging, spike density, glume color, glume hair, kernel color and vitreousness, respectively.

The analysis of variance of H’ for SPD, KCL and VTR showed significant ($P<0.05$) variation among geographical origins (Table 17), from which SPD and VTR showed significant ($P<0.05$) variation among genotypes with in geographical origins (Table 17).

Table 17. Mean squares and sum of squares for genotypes within geographical origins and between geographical origins from the ANOVA of H' for each qualitative traits.

Traits	Between geographical origins (df = 9)		Within geographical origins (df = 410)	
	Sum of square	Mean of square	Sum of square	Mean of square
Loading (LDG)	1.89	0.09 ns	33.86	0.09 ns
Spike density (SPD)	4.68	0.21 **	37.86	0.11 *
Glume color (GLC)	1.64	0.07 ns	35.53	0.09 ns
Glume hair (GLH)	0.95	0.04 ns	16	0.04 ns
Kernel color (KCL)	2.79	0.13 *	32.3	0.08 ns
Vitreousness (VTR)	4.86	0.22 **	41.41	0.11 *

df is degree freedom. ** and * are significant at p>0.01 and p>0.05, respectively. ns is non-significant at p>0.05.

5.2.3. Correspondence Analysis of Qualitative Traits

The pattern of diversity among studied durum wheat genotypes from different geographical origins where further analyzed by correspondence analysis of qualitative data. The computed inertia, percent of variance, percent of cumulative variance and coordinate values of geographical origins and qualitative traits are presented in Table 18. The first five components accounted for 61.6 % of the total variance and out of this the first two components explained 32.4 % of the total inertia, which means 32.4 % of the variability of the total genotypes, is explained by the plane demonstrated in Figure 5. The percentage of the two dimensions (32.4%) is greater than the reference value that equals 15.6 %, thus the variability explained by the two components is significant.

SPD is spike density; GLC is glume color; and KCL is kernel color. Ctr is contribution; and Dim is dimension Wilks test p-value is 2.964142e-09 which is Significant and indicates the variable factors are the best separated on the plane.

Figure 5. Distribution of the 10 geographical origins of the genotypes and qualitative traits along the first two axes of the components.

The studied qualitative traits with their three to four phenotypic classes discriminated the geographical origins very well. Dense spike and non-lodging phenotypic classes separate the cultivars from the remaining groups and slightly discriminate genotypes from North and East Shewa. SNNP and Tigray geographical origins were slightly differentiated by partly vitreous kernels and intermediate spike density. The remaining most geographical origins are placed at the center with vitreous kernel and low loading and hairiness traits.

Component 1 and 2 contained factors characterized by a positive and negative to the top and bottom and to the right and left axis of the graph, respectively. By phenotypic classes of qualitative traits, purple kernel color and black glume color strongly and positively contributed in component 1 while red-brown glume color contributed strongly and negatively. The strong positive contribution to component 2 was obtained from purple and brown kernel color phenotypic classes. While highly hairy glumes and dense spike phenotypic classes contributed strongly and negatively (Figure 5). In general, GLH, KCL and GLC traits highly contributed to component 1 and KCL, GLH and VTR traits highly contributed to component 2 (Table 18).

By geographical origins, the strong to weak positive coordinates on the axis of component 1 were shared among Bale-Harerge, Arsi, East Shewa, SNNP and North Shewa, and strong to weak negative contribution came from cultivars, Gojam, Tigray and Gondar-Wello (Table 18 and Figure 5). SNNP, Tigray, Bale-Harerge, Arsi, Gojam and Gondar-Wello were negatively contributed, and East Shewa, North Shewa and cultivars were positively contributed to the component 2.

Table 18. Eigen values, percent proportion of variance and cumulative variance of the first five coordinates, and coordinate scores of the 10 geographical origins and six quality traits.

	Dimensions				
	Co1	Co2	Co3	Co4	Co5
Eigenvalues	0.23	0.13	0.12	0.1	0.1
Proportion of variance (%)	20.4	12	11	9.6	8.4
Cumulative variance (%)	20.4	32.4	43.4	53.0	61.4
Geographical origins	Coordinate score of geographical origins				
Arsi	0.17	-0.06	0.03	-0.05	-0.14
Bale-Harerge	0.18	-0.07	0.01	0.00	0.04
East Shewa	0.10	0.08	0.04	0.02	-0.02
Gojam	-0.08	-0.05	0.00	-0.16	0.10
Gondar-Wello	-0.02	-0.02	0.07	-0.12	0.02
North Shewa	0.02	0.07	0.08	-0.01	0.04
SNNP	0.04	-0.16	0.10	0.03	0.24
Tigray	-0.04	-0.11	0.05	-0.05	0.03
Cultivars	-0.63	0.05	-0.47	0.36	-0.16
West Shewa	-0.04	0.02	-0.06	0.03	0.00
Qualitative traits	Coordinate score of qualitative traits				
Lodging (LDG)	-0.05	0.06	-0.07	0.01	0.05
Spike density (SPD)	-0.15	0.03	-0.10	0.13	-0.04
Glume color (GLC)	0.27	0.11	-0.01	0.10	0.01
Glume hairness GLH)	-0.80	0.26	1.17	0.09	-0.08
Kernel color (KCL)	0.58	0.35	0.15	0.34	0.03
Vitreousness (VTR)	-0.12	-0.18	0.01	0.26	0.88

Co1, Co2, Co3, Co4, and Co5 are coordinate 1, 2, 3, 4 and 5, respectively.

5.2.4. Comparison of the Performance of Selected Genotypes over all Individuals

Selection of top 5 % genotypes initially considers genotypes only with white kernel color and vitreous kernel. Then, priority is given for the quantitative traits GL and GPC contents. Because of the highly negative correlation between chemical quality traits (GL and GPC) with yield and yield-related traits, the selection considered the genotypes with better TGW and GY. The difference between the mean of the population and selected genotypes showed a significant t-test

value for GL and GPC as they are the criterion for the selection of genotype and were not significant for the remaining traits (Table 19).

Table 19. Comparison of the mean performances of selected top 5% genotypes ($\bar{x}$) with their corresponding mean performance of the population (μ).

Trait	Mean of the grand (μ)	Mean of the top 5% ($\bar{x}$)	Change of mean through selection (μ- $\bar{x}$)	% change of mean form population mean (μ)	t-test	
DH	71.4	70.0	1.4	2	-0.81	ns
DM	137.8	137.6	0.2	0.2	-0.23	ns
SPS	19.5	19.4	0.1	0.6	-0.21	ns
SPL (cm)	8.6	9.2	0.6	6	0.90	ns
PLH (cm)	108.7	108.7	0.0	0.0	0.00	ns
NET	6.7	6.8	0.1	1.8	0.16	ns
STD (%)	89.3	89.6	0.4	0.4	0.25	ns
TGW (g)	42.4	42.2	0.2	0.5	-0.07	ns
GY (t/ha)	8.4	8.8	0 .4	5	0.59	ns
GL (%)	25.7	30.5	4.8	15.7	3.78	***
GPC (%)	13.4	14.4	1	6.9	2.43	**

DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content. The t-value ($P<0.05$, df = 41) was 1.68. *** and ** are significant at $p<0.001$, $p<0.01$, respectively. ns is non-significant at $p<0.05$.

The mean GL of the top 5% genotypes was higher by 4.8 % over the total population mean (Table 19) with an average value of 30.5 % which meets the highest grain quality grade as described in Sissons (2016). These genotypes also showed 14.4 % GPC value that is above the optimal amount of GPC (>13 %) for the processing of durum wheat (Kadkol and Sissons, 2016). The Arsi zone contributed to most of the top 5% genotypes, followed by North Shewa, Bale-Harerge, and East Shewa. None of the cultivars were included in the top 5% of most performing genotypes due to their lower GL and GPC content.

Table 20. Top 5% performing genotypes (GEN) for overall mean of all quantitative traits as GL and GPC are criterion of selection.

Genotype	Geographical Origin	DH (day)	DM (day)	SPS	SPL (cm)	PLH (cm)	NET	STD (%)	TGW (g)	GY (t/ha)	GL (%)	GPC (%)
364	Bale-Hararge	67.8	135.3	20.2	9.0	106.3	9.0	91.3	40.4	7.88	35.4	15.5
150	North Shewa	68.3	136.5	19.1	8.9	109.8	5.9	91.3	39.4	8.66	34.8	15.0
191	East Shewa	69.5	137.3	19.0	8.4	108.5	5.3	90.0	38.0	8.42	33.8	15.6
286	Arsi	68.5	140.3	18.3	9.3	111.8	6.1	88.8	35.9	9.1	31.8	14.5
314	Arsi	67.3	140.5	18.6	8.1	106.5	7.2	85.0	44.0	8.17	30.9	14.2
288	Arsi	66.5	137.8	20.0	9.3	108.0	7.5	88.8	37.9	9.62	30.7	14.7
102	Bale-Hararge	71.3	136.0	20.9	9.2	107.0	6.1	91.3	38.9	8.56	30.6	14.3
124	East Shewa	73.3	135.5	19.6	8.4	112.8	5.7	87.5	34.1	8.46	30.3	14.1
44	Gondar-Wello	73.8	137.8	19.7	9.1	113.8	7.9	88.8	45.7	8.95	30.2	13.8
313	Arsi	67.0	136.0	19.1	8.2	98.3	9.4	87.5	43.4	8.18	30.2	13.9
386	Arsi	73.0	137.3	18.8	8.2	113.0	5.1	90.0	46.4	8.0	30.2	15.8
49	Gojam	67.5	138.0	17.3	8.5	107.8	6.9	86.3	41.4	7.7	29.5	13.6
244	North Shewa	74.8	137.8	20.6	7.7	107.0	7.3	87.5	38.1	6.56	28.7	14.0
245	North Shewa	68.5	135.3	19.6	10.1	114.3	6.9	90.0	49.2	10.1	29.4	14.4
292	Arsi	71.3	138.3	19.1	9.9	104.8	8.1	91.3	40.7	11.3	29.4	14.6
260	Bale-Hararge	74.3	139.0	20.3	9.7	107.0	6.1	88.8	51.6	8.31	29.2	13.8
284	Arsi	66.5	139.3	19.6	11.1	106.0	7.4	90.0	41.6	7.71	29.1	14.9
2	North Shewa	71.0	137.8	20.6	7.7	116.0	6.4	93.8	51.5	8.82	29.0	13.4
169	SNNP	72.8	137.5	20.1	9.8	114.3	6.2	91.3	40.7	10.6	29.0	14.2
193	Tigray	70.8	136.8	18.1	9.8	104.5	6.2	91.3	39.8	10.9	28.9	14.2
80	East Shewa	71.0	138.8	19.9	10.8	104.3	6.6	91.3	47.3	9.43	28.8	14.0
Mean		70.2	137.6	19.4	9.1	108.7	6.8	89.6	42.2	8.83	30.5	14.4

DH is days to heading; DM is days to maturity; SPS is number of spikelet per spike; SPL is spike length; PLH is plant height; NET is number of effective tiller; STD is stand growth; TGW is thousand grains weight; GY is grain yield; GL is gluten content and GPC is grain protein content.

5.3. Molecular Diversity Analysis of Ethiopian Durum Wheat using Microsatellite (SSR) Markers

5.3.1. Polymorphism of SSR Markers

In the present study, the molecular diversity of 104 Ethiopian durum wheat genotypes were evaluated using 14 microsatellite markers to screen the genotypes selected for several durum wheat quality traits required by end-user. The microsatellite primers used in this study revealed different ranges of amplified fragment size. All SSR markers showed a total of 119 alleles with an average of 8.5 alleles per locus. The number of alleles per locus (Na) ranged from 4 (Xwmc256) to 13 (Xwmc617). The gene diversity (h) ranged from 0.21 (Xwmc256) to 0.82 (Xwms294) (Table 21).

The highest major allele frequency (MAF) was obtained from locus Xwmc256, with the lowest gene diversity and polymorphic information content (PIC). In contrast, locus Xgwm294 showed the lowest major allele frequency, and the highest gene diversity as well as PIC. The PIC, which describes the extent of information retrieved from each locus, ranged from 0.25 (Xwmc256) to 0.8 (Xgwm294), with a mean value of 0.56 (Table 21).

Table 21. Diversity indices of the 14 simple sequence repeat (SSR) loci used in this study.

Number	Locus	Range of fragments (bp)	Na	MAF	h	PIC
1	Xwmc24	116-198	10	0.55	0.66	0.64
2	Xbarc240	256-398	6	0.61	0.56	0.50
3	Xgwm294	82-136	12	0.29	0.82	0.80
4	Xgwm47.1	132-210	11	0.55	0.66	0.64
5	Xbarc12	150-267	6	0.67	0.51	0.47
6	Xgwm493	138-224	10	0.62	0.56	0.52
7	Xbarc155	189-210	6	0.43	0.64	0.57
8	Xwmc617	190-262	13	0.35	0.77	0.74
9	Xgwm513	122-202	8	0.53	0.62	0.56
10	Xgwm120	116-196	12	0.40	0.77	0.75
11	Xgwm371	152-226	6	0.52	0.53	0.43
12	Xwmc256	106-152	4	0.88	0.21	0.25
13	Xbarc178	279-411	5	0.57	0.58	0.52
14	Xgwm46	144-224	10	0.63	0.58	0.55
		Mean	8.5	0.54	0.61	0.56

Na = Number of alleles per locus; h = gene diversity; MAF = major allelic frequency; and PIC = polymorphic information content.

5.3.2. Genetic Diversity within Populations

The diversity accompanied within and among populations is estimated by the distribution and frequency of alleles presented in that particular species. The distribution and frequency of alleles for a given locus in a given population determines the diversity of the population for that locus. The estimated numbers of allele per locus (Na) within a population ranged from 2.7 (West Gojam) to 4.2 (cultivars), with an average of 3.1 (Table 22). The number of effective alleles (Ne) ranged from 2.00 (West Gojam) to 3.22 (cultivars) with a mean value of 2.36. The Shannon diversity index ranged from 0.7 (West Gojam) to 1.17 (cultivars) with a mean value of 0.86. More than 60% of the populations scored an observed heterozygosity (Ho) value greater than the mean value of 0.29 which ranged from 0.24 (East Gojam) to 0.35 (cultivars). The percentage of polymorphic loci per population ranged from 81.5% for East Harerge to 100% for East Shewa and cultivars with an average value of 90% (Table 22).

West Shewa population appeared to have two private alleles for the locus Xwmc617 and Xgwm120. East Harerge, North Gondar, Arsi and cultivars showed single private allele at Xgwm46, Xgwm47.1, Xgwm493 and Xgwm47.1, respectively. From South Wello three private alleles were obtained from Xwmc24 (2) and Barc240, while East Shewa obtained two private alleles which came from Xgwm513 and Barc240.

Table 22. Mean diversity indices of 13 populations over 14 SSR loci.

	Population	N	Na	Ne	I	Ho	He	F	NPA	NCLA	%P
1	North Shewa	9	2.8	2.14	0.77	0.3	0.45	0.37	0.00	0.23	88.9
2	West Shewa	8	3.1	2.40	0.86	0.28	0.48	0.48	0.08	0.12	85.2
3	West Gojam	8	2.7	2.00	0.7	0.25	0.41	0.42	0.00	0.08	85.2
4	East Gojam	8	2.9	2.17	0.8	0.24	0.47	0.49	0.00	0.19	88.9
5	East Harerge	8	3.3	2.33	0.84	0.27	0.46	0.47	0.04	0.19	81.5
6	North Gondar	8	3.1	2.40	0.88	0.32	0.50	0.41	0.04	0.08	88.9
7	Bale	7	2.9	2.10	0.78	0.31	0.45	0.34	0.00	0.19	92.6
8	Arsi	7	3.1	2.47	0.9	0.32	0.51	0.36	0.04	0.00	88.9
9	South Wello	7	2.9	2.31	0.79	0.33	0.45	0.31	0.12	0.23	88.9
10	Central Tigray	7	2.9	2.29	0.82	0.3	0.47	0.38	0.00	0.00	92.6
11	Southern Tigray	7	3.1	2.20	0.78	0.29	0.43	0.32	0.00	0.00	88.9
12	Cultivars	10	4.2	3.22	1.17	0.35	0.62	0.48	0.04	0.04	100
13	East Shewa	10	3.7	2.63	1.04	0.28	0.59	0.56	0.08	0.08	100
	Mean	8	3.1	2.36	0.86	0.29	0.48	0.42	0.034	0.11	90
	SE		0.08	0.06	0.02	0.02	0.01	0.03			1.5

N is Size; Na is Number of alleles per locus; Ne is Number of effective alleles; I is Shannon diversity index; Ho is observed heterozygosity; He is expected heterozygosity; F is fixation index; NPA is number of private alleles; NCLA is number of common alleles, and %P is percent polymorphic loci.

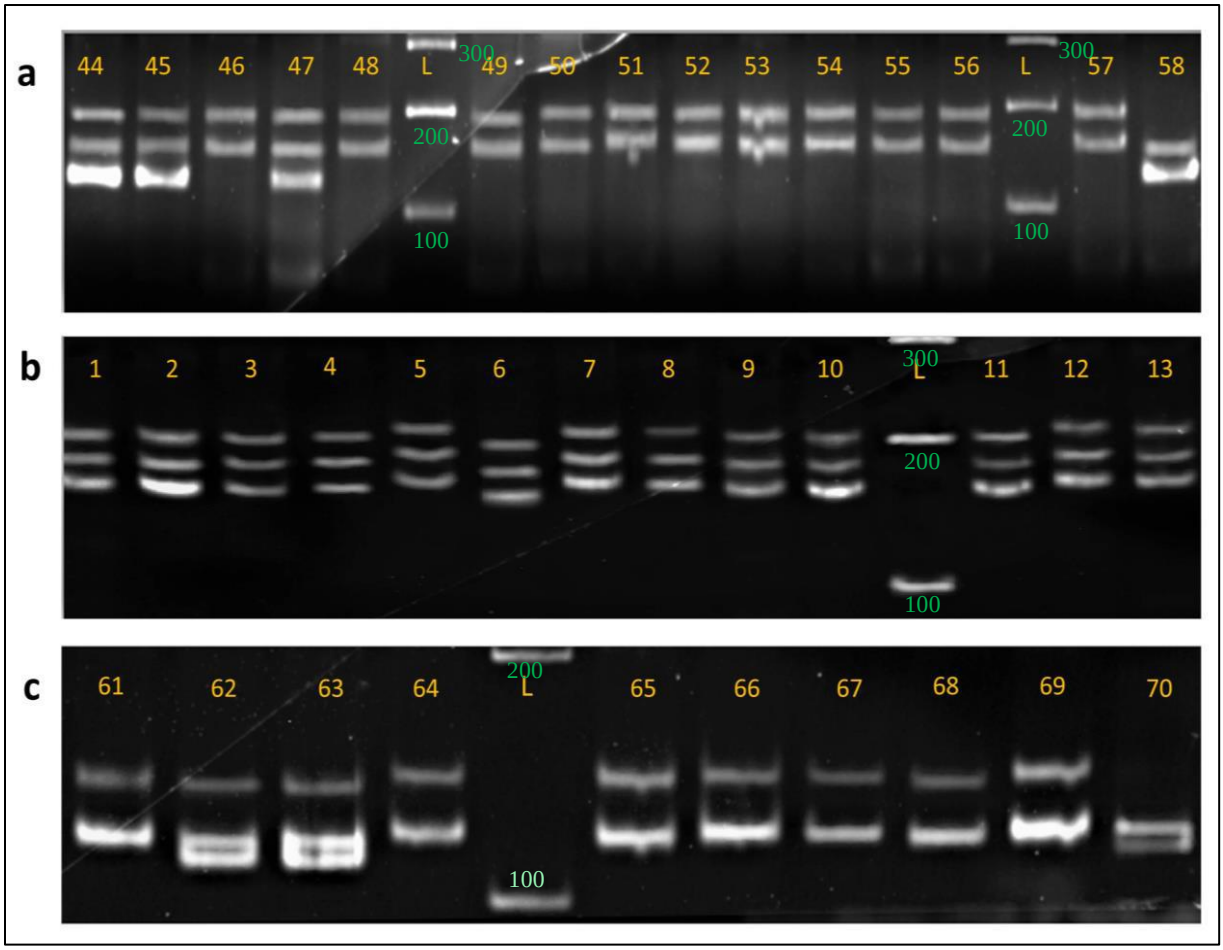


Figure 6. Electrophoresed gel image of SSR markers. (a) Primer Xwmc24; (b) Primer Xgwm46 and (c) Primer Xwmc256.

5.3.3. Genetic Relationship between Populations

The genetic relationship existing between populations was estimated by the pairwise population measure of Nei's genetic distance, and within and among population genetic differentiation. The pairwise Nei's genetic distance between populations was highest between cultivars and North Shewa, cultivars and East Harerge, and East Shewa and West Shewa (Figure 7). The second highest Nei's genetic distance was observed between Bale and cultivars, Bale and West Shewa, Arsi and East Harerge, and West Shewa and East Harerge (Figure 7). Generally, cultivars, West Shewa and East Harerge showed the higher pairwise genetic distance with other populations,

while South Wello, Central Tigray and South Tigray revealed lower pairwise genetic distance with other populations.

The pairwise population genetic differentiation was intermediate for most of the paired populations. The highest genetic differentiation was shown between Arsi and East Harerge, Arsi and East Gojam, Arsi and South Tigray, and cultivars and East Harerge. However, North Shewa had low pairwise genetic differentiation with most of the populations. Generally, Arsi and cultivars have shown high pairwise genetic differentiation value with most population. The genetic differentiation within populations was high in East Gojam, Arsi, Central Tigray, South Tigray, cultivars and East Shewa (Figure 7).

5.3.4. Analysis of Molecular Variance

The Analysis of molecular variance (AMOVA) revealed 88.35 % of the genetic variation found within populations and 11.65% of the genetic variation found among populations (Table 23).

Table 23. Analysis of molecular variance (AMOVA) shows the partition of genetic variation within and among populations of Ethiopian durum wheat genotypes.

	df	SS	MS	Estimate of variation	Percentage of variation	F-stat	P-value
Among populations	12	4.1	0.34	0.95	11.65%	FST = 0.12	0.001
Within populations	91	15.2	0.17	7.2	88.35%		
Total	103	19.2		8.15	100%		

df is degree freedom; SS is sum of squares; MS is mean square; FST is population differentiation.

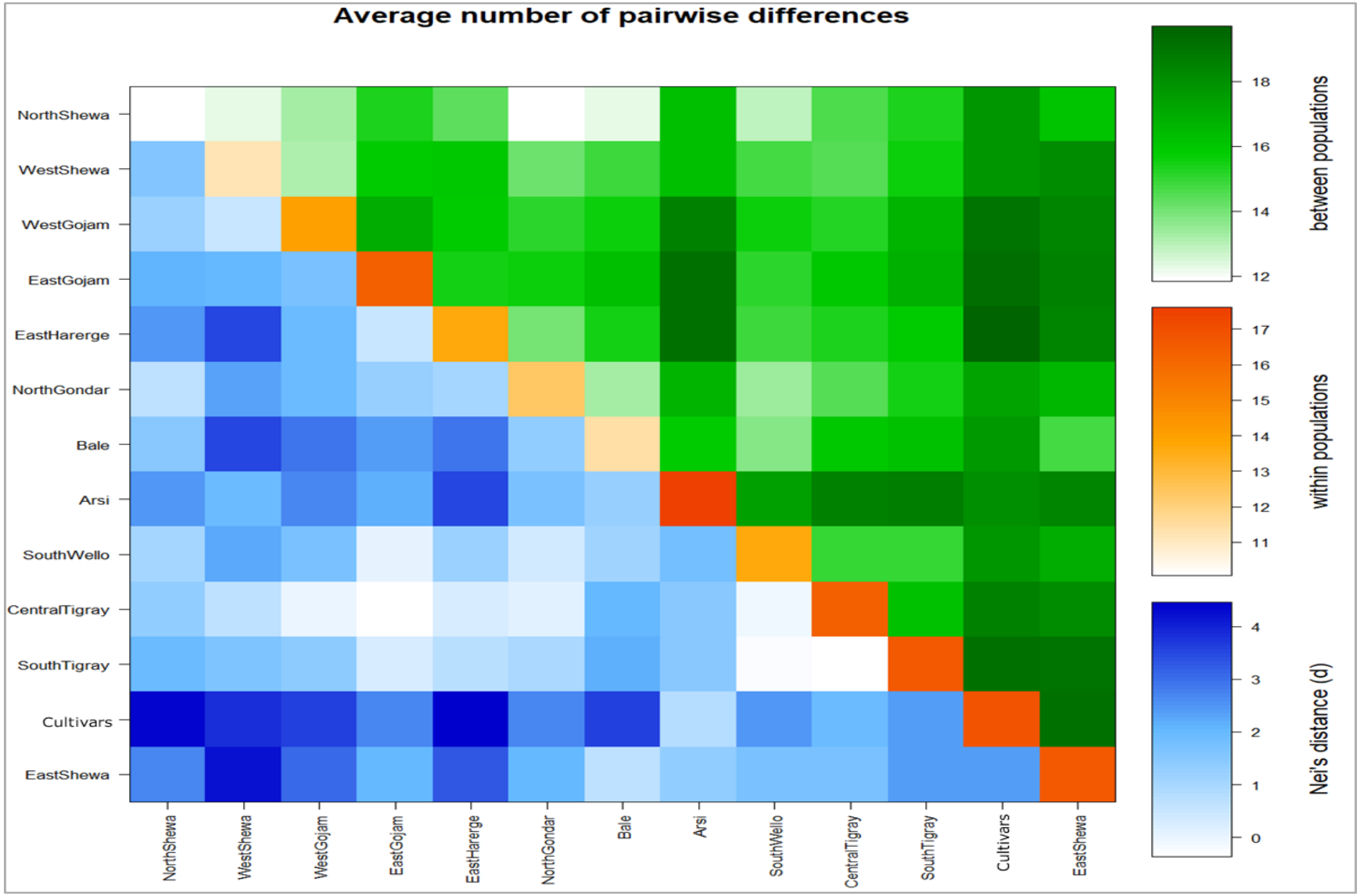
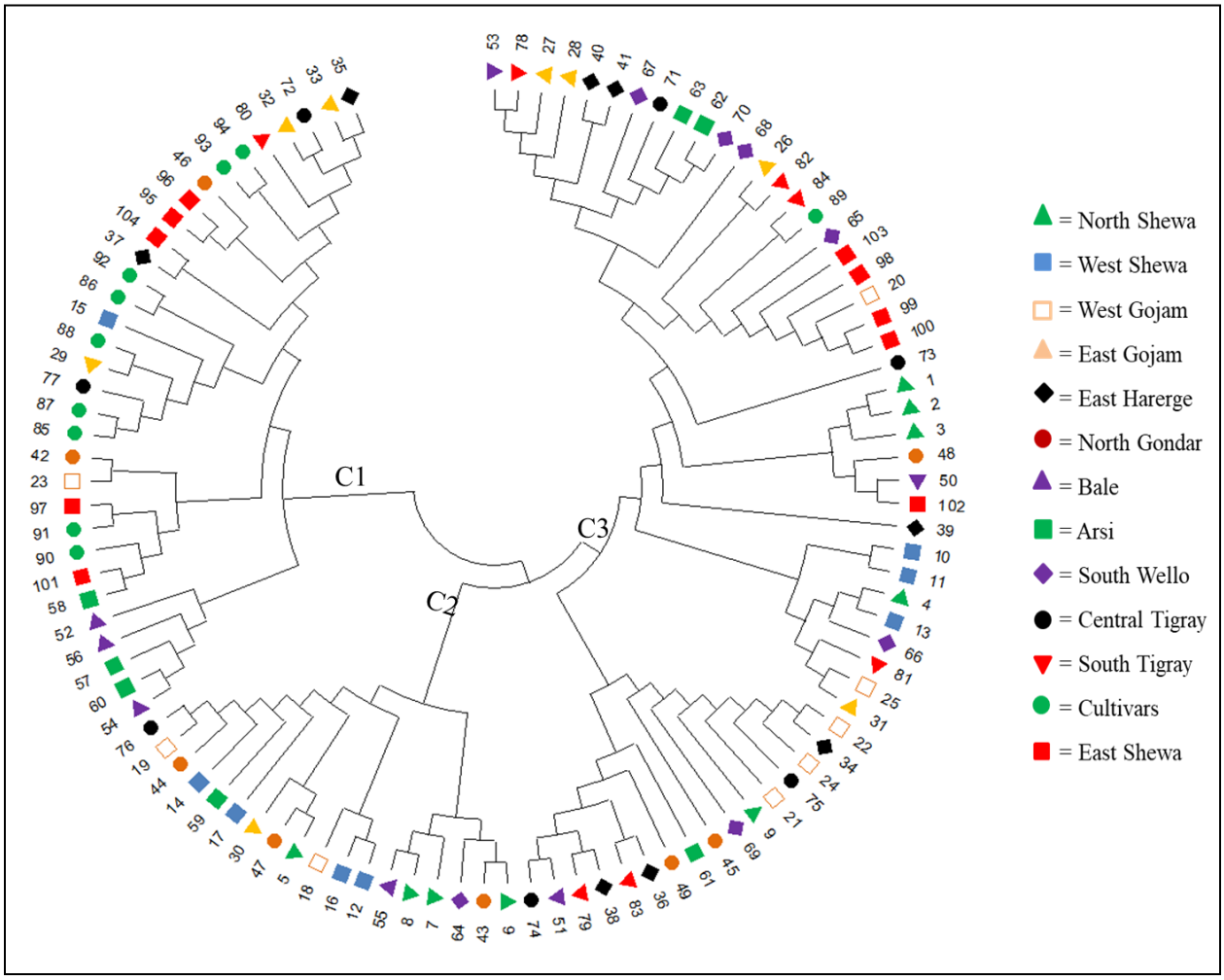


Figure 7. Nei's genetic distance and genetic differentiation within and among studied populations.

5.3.5. Cluster Analysis

The unweighted pair group method of Neighbor-Joining (NJ) based clustering of 104 Ethiopian durum wheat genotypes comprising the 13 populations' formed three major clusters (Figure 8).



C1, C2 and C3 are cluster one, two and three, respectively.

Figure 8. Neighbor-Joining dendrogram showing the pattern of genetic diversity among 104 Ethiopian durum wheat genotypes based on the 14 SSR markers.

Cluster 1 which holds 30.8 % of the studied genotypes was mostly composed of genotypes from cultivars and East Shewa populations (Table 24). Cluster 2 which holds 17.3% genotypes was mostly comprised from genotypes of North Shewa, West Shewa and North Gondar populations (Table 24). Member genotypes in the cluster 3 were comprised of all population and most

contribution came from South Wello, South Tigray and East Harerge. Generally, none of the clusters is comprised of a particular population.

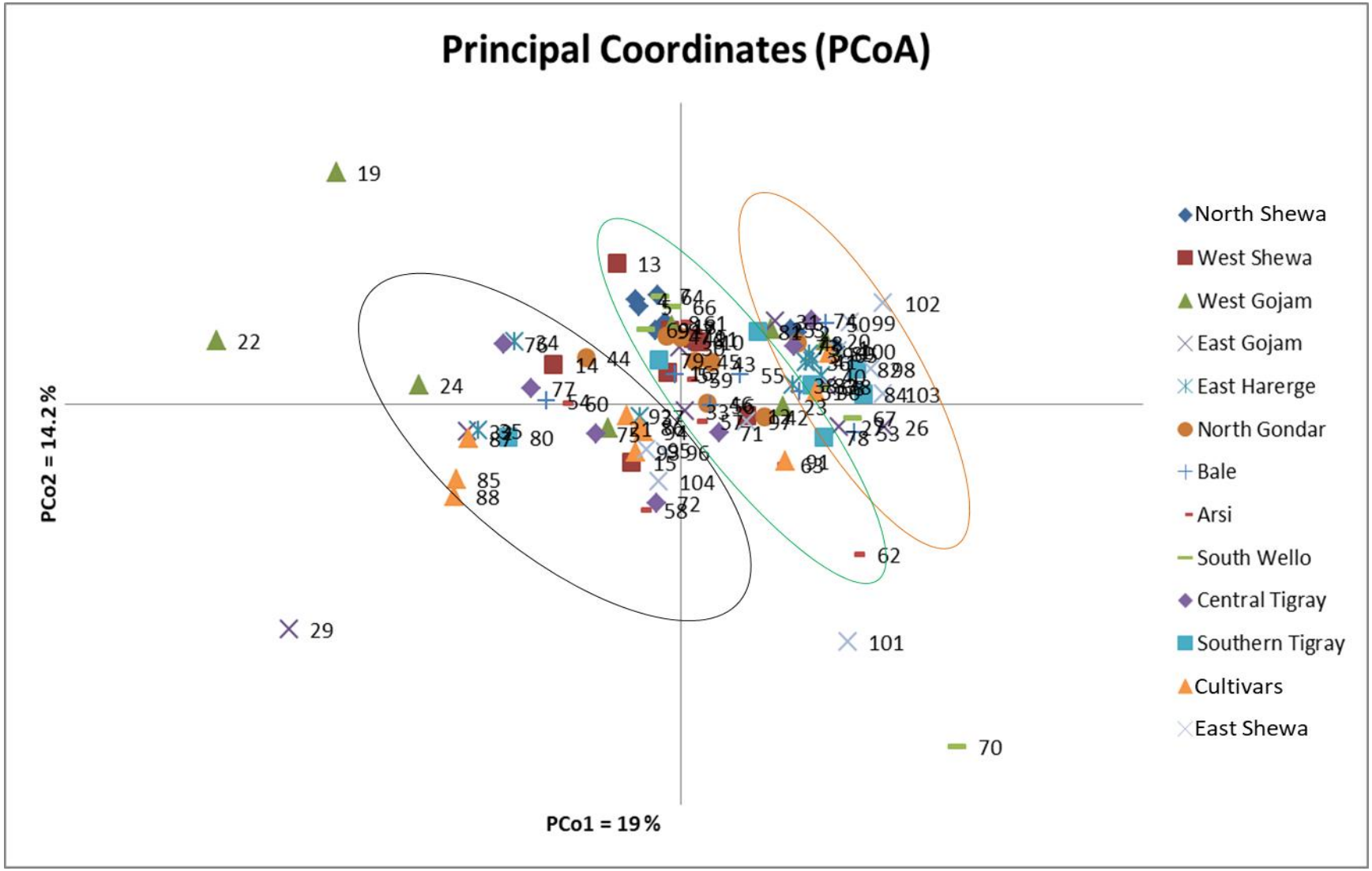
Table 24. Number of genotypes contributed by populations to the clusters.

Population	Size	Members of		
		C1	C2	C3
North Shewa	9	0	4	5
West Shewa	8	1	4	3
West Gojam	8	1	2	5
East Gojam	8	3	1	4
East Harerge	8	2	0	6
North Gondar	8	2	3	3
Bale	7	3	1	3
Arsi	7	3	1	3
South Wello	7	0	1	6
Central Tigray	7	2	1	4
Southern Tigray	7	1	0	6
Cultivars	10	9	0	1
East Shewa	10	5	0	5
Total	104	32	18	54

C1, C2 and C3 are cluster 1, 2 and 3, respectively.

5.3.6. Principal Coordinate Analysis

Principal coordinate analysis (PCoA) is a commonly used multivariate analysis to measure the extent of variance explained per principal coordinate and cumulatively. According to Quinn and Keough (2002), cluster analysis is more informative for closely related individuals, while PCoA is more sensitive regarding genetic distances among groups. The first three coordinates, PCo1 (19 %), Pco2 (14.2 %) and Pco3 (12 %) explained 45.2 % of the total variation and the first two principal coordinates grouped the genotypes into three main groups (Figure 9). This distribution and grouping of genotypes on the scatter plot of PCoA has confirmed the result of NJ clustering.

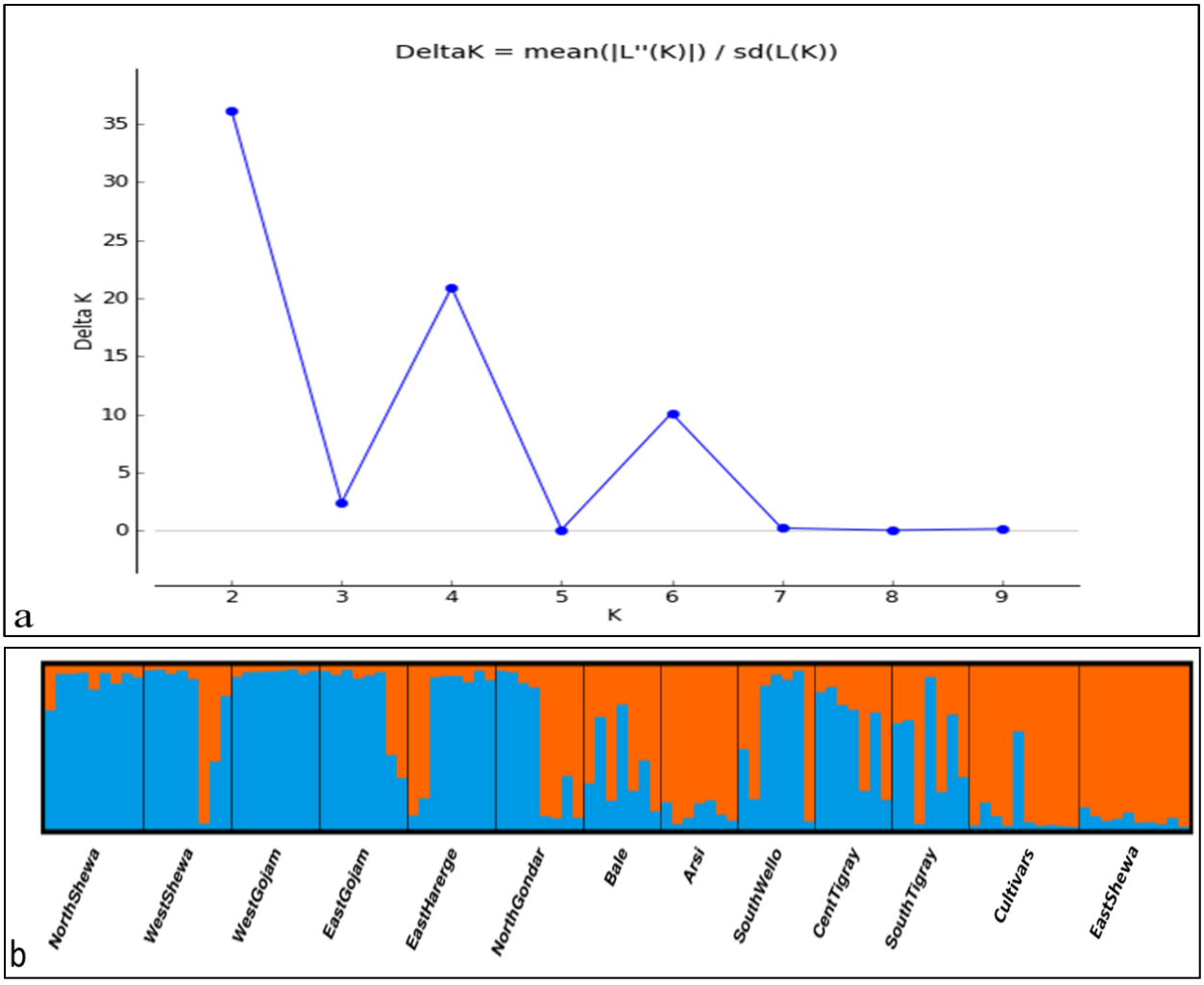


PCo1 and PCo2 are principal coordinate one and two, respectively.

Figure 9. Biplot of Principal Coordinate Analysis of 104 Ethiopian durum wheat genotypes.

5.3.7. Population Structure

The Bayesian model based analysis of population structure identified K = 2, as most likely number of cluster by reaching the maximum delta K values to a peak at K equals to 2 (Figure 10). The STRUCTURE bar graph (Figure 10) also provides information on the level of admixture in the studied samples. This identifies the intermixing of one genetic group with another genetic group. High admixture of genotypes was observed for North Gondar, Bale, Central Tigray and South Tigray.



(a) a biplot detected the best delta K value of 2 (the optimal number of clusters) based on Evanno *et al*. (2005) method; and (b) structure bar graph of populations at K = 2. The blue and orange colors represent the genetic groups found in the populations.

Figure 10. Population structure of 104 Ethiopian durum wheat genotypes.

6. Discussion

6.1. Phenotypic Diversity among Ethiopian Durum Wheat Genotypes Based on Quantitative Traits

6.1.1. Variations of Quantitative Traits among Genotypes

The observed wide range of values for most of the measured quantitative traits (Table 3) indicates the presence of high genetic variation among Ethiopian durum wheat genotypes. Similarly, wide phenotypic ranges in durum wheat genotypes reported by Khan *et al.*(2013), Dejene Mengistu and Pe (2016), Tesfaye Wolde *et al.* (2016a) and Taneva *et al.* (2019a) for GY, TGW and PLH; Hruskova and Svec (2009), Sayaslan *et al.* (2012) and Brankovic *et al.* (2018) for GPC; and Yagdi and Sozen (2009) for GY and GL.

Protein accumulation in grain could be affected by biotic and abiotic stresses. In cereals, many stress and defense proteins are accumulated in the grain to assist the massive synthesis of reverse protein and become activated during germination in response to environmental conditions (Huang *et al.*, 2011; Rocco *et al.*, 2019). The suitable GPC for processing of durum wheat to make pasta and other end products, 13% and above (Kadkol and Sissons, 2016), were scored by 285 genotypes from which 68 genotypes performed well at both locations implying that these 68 genotypes were not significantly affected by the environmental conditions of the test locations and may have broader adaptation for GPC.

The gluten content in grains of durum wheat is the determinant of the gluten content in the flour. Thus, GL of the grain is crucial in processing or pasta making because gluten confer elasticity and viscosity to the dough (Padalino *et al.*, 2019). The 97 genotypes, which scored above 28 % of GL with the highest from 364, 150, 191, 287, 354, 263, 282, 285, 294 and 67, meets the

highest grain processing quality grade as described in Sissons (2016). All cultivars scored below 28 % for GL because yield potential was mainly targeted during their improvement. Although the genotype by location interaction was significant for GL, 24 landraces scored above 28% at both locations, which implies that these landraces may be resistant to environmental stresses and have the adaptive potential for ranges of environments.

The highly significant ($P<0.01$) variations among genotypes for most of the studied traits as identified by ANOVA (Table 5) indicate the presence of an adequate amount of genetic variability among characterized genotypes. These genotypic variations generally might be caused by hybridization, mutation, recombination and/or related phenomenon. Similarly, Dejene Mengistu and Pe (2016) found a significant difference among 49 Ethiopian durum wheat genotypes (comprised of 45 landraces and 4 cultivars) for DM, PLH, SPL, TGW and GY, and non-significant for NET. Dejene Mengistu *et al.* (2016) and Tesfaye Wolde *et al.* (2016a) reported significant variation among 289 and 68, respectively, Ethiopian durum wheat genotypes for DM, PLH, NET, SPL, SPS, TGW, and GY. Yohannes Alemu *et al.* (2020) also reported a significant difference between 64 Ethiopian durum wheat landraces for traits DH, DM, PLH, SPS, SPL, TGW, GY and GPC. Non-Ethiopian durum wheat collections have also showed significant variation among genotypes for GL, GPC, GY, TGW, SPS, SPL and PLH (Khan *et al.*, 2013; Sourour *et al.*, 2018; Johnson *et al.*, 2019). This indicates that durum has a wider adaptation in different agro-ecological conditions.

The highly significant ($P<0.01$) genotype by location interaction for all of the quantitative traits except STD (Table 5) resulted in this study and significant genotype by location interaction reported in previous studies for GPC, TGW and GY (Bilgin *et al.*, 2010; Johnson *et al.*, 2019; Rapp *et al.*, 2019; Taneva *et al.*, 2019b) implies the performance of Ethiopian durum wheat

genotypes could be affected significantly by environmental conditions and indicates that some genotypes perform best at specific locations. In contrast, Dejene Mengistu *et al.* (2016) found non-significant genotype by location interaction for all studied traits of their study (DM, NET, PLH, SPL, SPS, GY and TGW).

The observed closeness between the values of GCV and PCV for DH, DM, SPS, SPL, PLH, TGW, GL and GPC indicates the environmental influence on the phenotypic expression of respective traits is small/less. While, the biggest difference in GY, STD and NET suggests the presence of higher environmental influence than the genotypic factor for the variation. Similarly, close difference between the values of GCV and PCV reported by Mohammed Abinsa *et al.* (2012) for TGW and GPC, and Tesfaye Wolde *et al.* (2016a), Yohannes Alemu *et al.* (2020) and Bilgin *et al.* (2011) for DH, DM, SPS, SPL, PLH, and TGW. In addition, wide difference between the values of GCV and PCV for GY also reported by Yohannes Alemu *et al.* (2020) and Bilgin *et al.* (2011) indicates yield traits are significantly affected by environmental factors.

6.1.2. Selection Efficiency of Quantitative Traits in Durum Wheat

The knowledge on heritability and genetic advance enable to make parental selection wisely (Tuhina-Khatun *et al*, 2007), predict the nature of succeeding generations, and estimate the magnitude of genetic improvement through selection (Yadav *et al.*, 2015). Assuming 5% selection intensity, high to moderate heritability pulled with high to moderate expected GAM indicates the additive genes effect while high heritability coupled with low genetic advance reveals the non-addictive gene effects for a given trait. In this study, the additive genes effect was found for SPL, TGW and GL which indicates the possibility of undertaking an effective selection of genotype through phenotype expression at early generation and enables to achieve effective breeding progress on improvement of these yield and quality traits in durum wheat.

Similarly, additive genes effect for TGW has been reported by Mohammed Abinsa *et al.* (2012), Dejene Mengistu *et al.* (2016), Tesfaye Wolde *et al.* (2016a), Taneva *et al.* (2019b), and Yohannes Alemu *et al.* (2020). But, Bilgin *et al.* (2010), Bilgin *et al.* (2011), Dejene Mengistu and Pe (2016), and Dejene Mengistu *et al.* (2016), found non-additive genes effect for TGW due to low genetic advance. In the present study, medium heritability with moderate genetic advance was found for GY (49.4 % and 17.4, respectively), indicating the influence of the non-additive genes effect. Similarly, the non-additive genes effect of GY was reported by Tesfaye Wolde *et al.* (2016a) and Getachew Belay *et al.* (1993). Therefore, selection for GY in durum wheat would be extended to generations.

Morphological and phenological traits appear to have high heritability values, as these traits are known to be highly influenced by genetic factors than environmental factors. Early heading is desirable in wheat to broaden the time for grain filling period that maximize the weight of grains. In addition, early growing plants are more preferable in drought areas and changing climate conditions. In line to our finding, the high value of heritability for DH had been reported by Tesfaye Wolde *et al*. (2016a) (87%), Dejene Mengistu and Pe (2016) (72%), Yohannes Alemu *et al*. (2020) (76.7%), and Khan *et al*. (2013) (53%). The observed non-additive genes effect of DH indicates the effective selection of the trait would be extended to generations. Similarly, Getachew Belay *et al.* (1993), Dejene Mengistu and Pe (2016) and Yohannes Alemu *et al.* (2020) found non-additive genes effect of DH, while additive genes effect reported by Dejene Mengistu *et al.* (2016) and Tesfaye Wolde *et al.* (2016a).

The significant influence of environment on the expression of protein traits minimizes the heritability value of both GL (55.6 %) and GPC (44.8 %) (Table 6). The GPC showed moderate heritability with low GAM value, thus it should be considered that the selection for GPC in early

and advanced generations could be complicated. Similarly, moderate heritability with low genetic advance value for GPC were reported by Akcura *et al*. (2009), Bilgin *et al*. (2010), Mohammed Abinsa *et al.* (2012) and Yohannes Alemu *et al*. (2020). The high heritability with a moderate GAM value of GL indicates the presence of the additive genes effect of the trait. This suggests that the possibility of undertaking effective selection of genotype through phenotype expression at early generation and enables to achieve effective breeding progress on improvement of GL in durum wheat. Previously there were limited numbers of researches conducted to determine the heritability value of GL in durum wheat. Similarly, Brankovic *et al*. (2018) reported high heritability and moderate GA for wet gluten content. In addition, Mohammed Abinsa *et al.* (2012) and Taneva *et al*. (2019b) reported moderate heritability with moderate and low GA, respectively, for wet gluten content.

6.1.3. Associations of Quantitative Traits in Durum Wheat

Understanding the association of different traits with the target trait is the most important step in the breeding program for the improvement of a target trait using an indirect selection of other traits. Hence, a correlation coefficient helps in quantifying and identifying the magnitude and the direction of the influence of component traits on the main traits.

Proteins with other traits

The GPC of studied genotypes showed a significant negative phenotypic correlation with most yield and yield-related traits. The identified negative and highly significant ($P<0.001$) correlation between GPC and GY (-0.22) is in parallel with Bilgin *et al*. (2010) (-0.29), Gulmezoglu *et al*. (2010) (-0.46), Taneva *et al*. (2019a) (-0.47) and Yohannes Alemu *et al*. (2020) (-0.21). However, Grahhman *et al*. (2014) reported a positive and highly significant correlation of GPC

and GY (0.6) which deviate from the commonly seen negative correlation of the two traits. The observed negative and significant (P<0.05) correlation between GPC and TGW (-0.11) has also been reported in Wang and Fu (2020) (-0.94), Grahmann *et al.* (2014) (-0.53), Taneva *et al.* (2019a) (-0.48) and Blanco *et al.* (2012) (-0.68).

Limited numbers of researches are dealt with the GL of durum wheat and mainly studied for gluten index, and wet and dry gluten content. Sourour *et al*, (2018) reported a positive and highly significant correlation of GL with PLH (0.44), SPL (0.39), SPS (0.47), and GPC (0.94) while negatively correlated with GY (-0.32). This report is totally in agreement with our findings except for the positive correlation of GL with SPS. Comparably, Yagdi and Sozen (2009) found a negative and significant correlation of GL with STD (-0.27) and SPS (-0.27). In terms of wet gluten content, Bilgin *et al*., (2010) identified a positive and significant correlation with GPC (0.71), however, negatively and significantly correlated with TGW (-0.68) and GY (-0.32).

Generally, GL and GPC showed a negative and significant correlation with most yield and/or yield-related traits (Table 7). Hence, breeding for the improvement of these processing quality traits will decrease the yield of durum wheat. Due to that, selecting genotypes which have high GPC and GL along with high TGW and GY is more potentially important to be considered for processing or as raw material for further improvement through breeding.

Grain yield with other traits

As GY is the output gained from the summation of the weight of all grains from the cultivated crop, the weight of total grain harvested is expected to be correlated positively with the weight of one thousand seeds. In line with our finding, previous studies showed a positive and highly significant (P<0.001) correlation between GY and TGW (Bilgin *et al*., 2010 (0.34); Ali, 2012

(0.85); Khan *et al*., 2013 (0.35); Dejene Mengistu *et al.,* 2016 (0.48)). This indicates the improvement of either GY or TGW will improve the other trait in durum wheat.

The strong negative correlation of GY with PLH (-0.34) was also confirmed by Khan *et al*. (2013) (-0.31) and Sourour *et al*. (2018) (-0.19), which indicates the improvement of either GY or PLH will decrease the performance of the other trait in durum wheat. However, Dejene Mengistu *et al*. (2016), Yagdi and Sozen (2009) and Bilgin *et al*. (2011) reported a positive and highly significant ($P<0.001$) correlation between GY and PLH (0.36, 0.40, and 0.22, respectively). The identified negative and highly significant phenotypic correlation between GY and DH (-0.31) is in line with Bilgin *et al*. (2011) (-0.27), Dejene Mengistu *et al*. (2016) (-0.31), Tesfaye Wolde *et al*. (2016b) (-0.18), and Yohannes Alemu *et al*. (2020) (-0.24). This negative association between GY and DH may be due to the reduction of water availability as the days of heading extended which decrease the yield of the crop.

6.1.4. Clustering and Patterns of Distribution

The results of PCA and cluster analysis strongly support each other. Geographical origins clustered together in Figure 3 were also placed near each other in the same quadrant except for cultivars and Arsi in the PCA analysis in Figure 4. The 11 quantitative traits classified the geographical origin well, with TGW and GY separated cultivars from the remaining landraces, and GPC and GL discriminated some landraces from Arsi. This suggests that different traits have different contributions to the variation explained by each PC. However, genotypes collected from Gondar-Wello, SNNP, Tigray, Bale-Harerge, and Gojam landraces were placed near each other. This revealed Ethiopian durum wheat landraces collected from such geographical regions shared a common gene pool due to the possibility of germplasm exchange during socio-economic interactions, gene flow, and/or duplications of germplasms in the landrace collections. For

instance, the clustering of SNNP and Tigray landraces was close and it might be due to the relocation of farmers together with the landraces because of drought, conflict, or some other cases. In contrast to our finding, Fassil Kebebew *et al*. (2001), Firdissa Eticha *et al*. (2005) and Faris Hailu *et al*., (2010) clustered landraces from the same sites of origin into different multiple clusters. This may be due to the construction of the populations from larger areas instead of smaller areas collections, for example in Fassil Kebebew *et al*. (2001). Lumping of species together from broader areas has effects on the clustering pattern of that species and may also bias estimates of diversity (Seifu Tsegaye *et al*., 1996).

For better genetic recombination, the crossing of genotypes included in clusters with larger inter-cluster distance may improve the adaptability and genetic constituent of Ethiopian durum wheat germplasm and also help in the successful selection of genetically divergent parents for breeding programs (Allard, 1960; Arega Gashaw *et al*., 2007; Singh, 2007). The most divergent cluster group in this study was C4 and C5 (D^2 = 58.7) which constituted of Arsi (mainly) and cultivars, respectively, populations. Relatively high values of inter-cluster distances may result due to their differences shown for GL and GPC (higher in Arsi) and yield traits (higher in cultivars) by the genotypes included in the clusters. Hence, crossing 263, 197, 277, 150 and 276 (best genotypes of cluster 4) with 416, 397, 415, 413 and 402 (best genotypes of cluster 5) (Table 11) may result in successful offspring with high GL, GPC and yield-related traits. The lowered inter-cluster distances observed between some clusters indicates that members of these clusters were closely related and, therefore, the crossing of genotypes from these pair of clusters may not produce a good level of heterotic expression and variability in early generations (Allard, 1960).

Characters with larger absolute values closer to one within the PC influence the clustering more than those with lower absolute values closer to zero (Chahal and Gosal, 2002). Therefore,

phenotypic traits with high coordinate values are the most discriminant traits for those PCs, however, differentiation of genotypes into different clusters is governed by the cumulative effects of traits. Rencher (2002) explained that the first PC is the linear combination with maximal variance; the second PC is the linear combination with a maximal variance in a direction orthogonal to the first PC. Accordingly, DH, SPS, STD, GL and GPC had high contributions to the two PCs (Figure 4). The placement of GL, GPC and TGW on 2D plane of the first two PCs is similar to Fiore *et al*. (2019). Dejene Mengistu *et al*. (2015) also reported a similar arrangement of DM and DH on the 2D plane of the first two PCs. In contrast, Sourour *et al*. (2018) who analyzed ten durum wheat genotypes reported different contributions of GY, TGW, PLH, SPL, SPS, GL and GPC on the first two PCs.

6.2. Phenotypic Diversity among Ethiopian Durum Wheat Genotypes Based on Qualitative Traits

6.2.1. Distribution and Variability of Qualitative Traits

Analyzed durum wheat genotypes were polymorphic within studied three to four phenotypic classes of qualitative phenotypic traits (Table 14). The Existence of technologically important phenotypic classes in high frequency in Ethiopian durum wheat genotypes gives an opportunity to select valuable genotypes for the processing sectors. Among these traits, white kernel color and vitreous kernel were the dominating phenotypic classes. Similarly, Faris Hailu *et al*. (2010) reported the dominance of white kernel color in Ethiopian tetraploid wheat. However, Dejene Mengistu *et al*. (2015) reported the rarity of white kernel color in Ethiopian durum wheat landraces. The observed purple kernel color in this study was also confirmed by Efrem Bechere *et al*. (1996a) who found 16 of % genotypes with purple-colored kernels in their collection.

Purple-colored durum wheat kernels are endemic to Ethiopia (Zavan *et al*., 1991) and classified as *Triticum turgidum sub. aethopicum* (Firdissa Eticha *et al*., 2005).

The presence of hairy glumes in the collections of Gondar-Wello, North Shewa is similar to Efrem Bechere *et al*. (1996a). Mulugeta Negassa *et al*. (1986) suggest that the hairy glume phenotypic class of durum wheat in Ethiopia may be associated with powdery mildew resistance. The observed high dominance of hairless subclass of GLH is in line with Efrem Bechere *et al*., (1996a), Seifu Tsegaye *et al*. (1996), and Firdissa Eticha *et al*. (2005). In contrast, a high level of polymorphism for GLH in Ethiopian durum wheat was reported by Endashaw Bekele (1984).

Genotypes with dense spike type were rare in Arsi, Bale-Harerge, Gondar-Wello, SNNP and Gojam, while predominant in cultivars and less common in West Shewa. This finding was in contrast to Faris Hailu *et al*. (2010) who reported the dominance of the dense spike in most parts of Ethiopia except Bale and Wello. Dejene Mengistu *et al*. (2015) and Efrem Bechere *et al*. (1996a) also reported the predominance of dense spike in all Ethiopia regions. According to Firdissa Eticha *et al*. (2005) and Seifu Tsegaye *et al*. (1996), lax spike is associated with resistance to spike-related diseases. Black glume color was rare in all geographic origins as the finding of Jain *et al.* (1975). Mulugeta Negassa *et al*. (1986) also showed the association of black glume color with gluten quality.

Chi-Square was used to test the difference between what is actually observed in the data and what would be expected if there was no variation. The observed significant difference in chi-square value revealed the existence of variability within Ethiopian durum wheat genotypes for all qualitative traits considered. The significance of location for LDG and GLC over combined test sites identifies growing environment had significant effects on the performance of durum wheat

genotypes for these traits. While location had no significant effect on the remaining qualitative traits, implies these traits are more governed by genetic than environmental factors.

The overall mean of Shannon diversity index value was low (H'<0.40) for all traits except LDG and that indicates the presence of low genetic diversity in Ethiopian durum wheat genotypes for most of the studied qualitative traits. The low H' value recorded for GLH in this study was also similar to Dejene Mengistu *et al*. (2015) indicates the dominance of a given phenotypic class of GLH in Ethiopian collections. The observed significant (P<0.05) variation of SPD, VTR and KCL among geographic origins was contradicted with Firdissa Eticha *et al*. (2005) who reported a non-significant variation of SPD and KCL among geographical origins. The identified significant (P<0.05) variation of SPD and VTR among genotypes within geographic origins was similar to Dejene Mengistu *et al*. (2015). This confirms the presence considerable genetic variation among Ethiopian durum wheat genotypes for SPD, KCL, and VTR traits.

The placement of most geographical origins near to each other at the center of the two components (Figure 5) indicates the presence of high genetic similarity may rise due to high gene flow through seed exchange during marketing and relocation of farmers with their seed. In addition the separation of cultivars from the landraces indicates the presence of genetic variation by having the rare traits in the landraces like dense spike and no lodging.

6.2.2. Promising Landraces for Processing Sectors and Breeding Programs

Genotypes included in the top 5% were totally landraces (Table 20) and none of the cultivars were included due to their lower GL and GPC since they were highly selected for yield traits. The mean of the top 5% genotypes showed deviation from the population mean for GL and GPC by showing higher value as they are the criterion for the selection of genotype. However, traits

correlated negatively (e.g., GY and TGW) with the protein traits were able to be maintained by the top 5% genotypes which suggests these landraces with optimal GL and GPC along with sufficient yield traits should have to be directly assessed for technological traits variations to find the very best landrace for the production of quality pasta and related products and/or in wheat breeding programs to improve both protein and yield related traits.

6.3. Microsatellite Diversity of Ethiopian Durum Wheat Genotypes

6.3.1. Diversity of Microsatellite Markers

Many studies have shown that microsatellite markers are the best molecular marker for genetic diversity studies and fingerprinting in crop plants (Roder 1998; Mun *et al*., 2006; Kumar *et al*., 2009; Achtar *et al*., 2010). Durum wheat is an allotetraploid species whose amplification with SSR markers may resulted in up to 4 alleles per individual at a locus.

According to Vaiman *et al.* (1994), 11 out of 14 loci were highly informative (PIC ≥ 0.5) and the remaining loci (Xbarc12, Xgwm371 and Xwmc256) were moderately informative (0.25≤PIC<0.5). The frequency of major alleles across each locus was less than 0.67 for all SSR markers except Xwmc256 (0.88) coupled with the higher value of PIC, which indicates that all markers were polymorphic and signifying the usefulness of SSR molecular tools for genetic diversity studies. The identified 8.5 mean number of alleles per locus for all markers were similar to Mondini *et al*. (2010), who analyzed 23 Ethiopian durum wheat landraces using 28 SSR markers, found an average of 8.7 alleles per locus and 0.61 mean PIC. However, lower than the report of Yifru Teklu *et al*. (2006) who analyzed 12 durum wheat genotypes using 29 SSR markers and found 11.03 alleles per locus. In addition, Sintayehu Alamerew *et al*. (2004) found 7.9 numbers of alleles per locus for 22 SSR markers among 12 durum wheat genotypes. The

identification of many alleles at a given microsatellite locus in previous studies and the current study identifies the existence of high genetic diversity in Ethiopian durum wheat genotypes.

In non-Ethiopian collections, Mangini *et al*. (2010) presented 9.45 mean Na and 0.71 mean PIC values for 11 SSR markers applied on 28 durum wheat cultivars. In 37 Syrian durum wheat genotypes, Achtar *et al*. (2010) reported an average of 8 Na with a mean PIC value of 0.57 for 34 loci. In Italian 164 durum wheat genotypes, Marzario *et al*. (2018) found 4.1 mean Na across 44 SSR markers. From Tunisia, Medini *et al*. (2005) reported 10.4 alleles per locus with an average 0.72 PIC value for 15 SSR markers among 24 durum wheat cultivars. In addition, Ouaja *et al*. (2021) who analyzed 304 Tunisian durum wheat landraces using 10 SSR markers found that a mean of 9.9 alleles per locus with a mean PIC value of 0.69. Syrian and Moroccan durum wheat were analyzed by Kehel *et al*. (2013) and found 16 numbers of alleles per locus for 51 SSR markers. This indicates that durum wheat has high genetic diversity worldwide.

The three microsatellites, in this study, Xgwm294, Xwmc617 and Xgwm120 have showed 6 to 7 number of alleles per locus with higher h and PIC value, which indicates the high discriminating efficiency of these microsatellites for genetic diversity study of durum wheat genotypes.

6.3.2. Population Genetic Diversity

The high mean number of alleles per locus (3.1), effective number of alleles (2.36), Shannon diversity index (0.86), and percent of polymorphic loci (90%) observed across the 13 populations revealed the existence of high genetic diversity within Ethiopian durum wheat populations. The high genetic diversity of durum wheat populations had been reported previously in various parts of the world. In Italy, Marzario *et al*. (2018) reported a mean Na value of 3.09 and an expected heterozygosity value of 0.53. Kehel *et al*. (2013) also reported the presence of a high mean

number of alleles in Moroccan and Syrian durum wheat landraces. In contrast, lower genetic diversity in Tunisian durum wheat was reported by Ouaja *et al*. (2021) who identified very low observed heterozygosity (0.075) and high fixation index (0.85).

Previously, the genetic diversity of Ethiopian durum wheat genotypes was reported by many scientists. Among these, Mondini *et al*. (2010) reported a high percentage of polymorphism (91%), 4.87 numbers of alleles per locus and low observed heterozygosity (0.14) across 9 populations. Sinatyehu Alamerew *et al*. (2004) also reported the presence of 24 durum-specific alleles among 12 Ethiopian durum wheat genotypes which reflect the existence of the diverse population and rich genetic diversity. Comparatively, among the studied populations, the cultivars had the highest values for all genetic diversity indices as they might have developed from different genotypes.

6.3.3. Genetic Relationship and Population Structure

The genetic relatedness of genotypes was identified by the NJ method of clustering and PCoA using 14 SSR markers which were used to describe the relationship between genotypes. Interestingly, as presented in Figure 8 all cultivars except one were clustered in cluster 1 with some landraces. This common morphological trait may be due to lumping of cultivars with landraces and/or the cultivars may have been developed from the local landraces. Similarly, Slim *et al*. (2020) grouped Tunisian 41 landraces and 13 varieties into 2 clusters in which varieties were clustered together with landraces. Arora *et al*. (2014) also clustered durum varieties together with landraces. This indicates that landraces and cultivars share similar genetic background. In contrast, Medini *et al*. (2005) worked on 25 landraces and 9 varieties, were clustered the varieties in one group distinctively from the landraces. Cultivars showed clear separation from landraces using SSR markers related to Ethiopian stem rust of durum wheat in

the report of Jemanesh Kifetew *et al*. (2013a). The clear separation might be due to differences in adaptation of durum wheat to a certain disease related traits. The PCoA which showed 3 groups of genotypes on the 2D scatterplot confirm the result obtained by the cluster analysis (Figure 9). The placement of some cultivars with landraces identifies the relatedness of those cultivars with landraces as they share similar morphological appearances and genetic backgrounds. In other cases, landraces were not usually considered in the breeding program due to their tall stature, lateness, and disease susceptibility (Kirouani *et al*., 2018).

The moderate genetic variation among populations as revealed by AMOVA indicates the presence of lower genetic variation between studied populations. Similar moderate genetic variation among populations was reported in Ethiopia by Mondini *et al*. (2010) (18.76%), in Tunisia by Ouaja *et al.* (2021) (19%) and Slim *et al.* (2020) (22%), in Mediterranean regions by Amallah *et al.* (2016) (16%), and in India by Arora *et al.* (2014) (21.3%). The high genetic variation within the population observed in this study could be attributed to mutation, recombination, hybridization and/or related phenomenon. The lower proportion of genetic variation among populations (11.35%) compared to within populations (88.65%) may be due to durum wheat grains exchange through marketing and relocation of farmers with their seed.

The population structure analysis identified that some populations of landraces were not found to be distinct and many landraces received alleles from two gene pools (Figure 10). Although studied genotypes showed similar morphological appearance in terms of white kernel color, vitreous kernel, high GL and GPC, the population were structured into two genetic groups. This grouping may be due to the relatedness of genotypes for the trait associated with the SSR marker, sharing of similar genetic background and/or high gene flow. Hence the grouping is

based on the presence of genetic differences among genotypes at the investigated loci which may be caused by mutation, translocation, recombination, and some other phenomenon.

Similarly, Oliveira *et al*. (2012) among 174 durum wheat genotypes from the Mediterranean basin and Kehel *et al*. (2013) among Moroccan (98) and Syrian (90) durum wheat landraces reported 2 genetic groups with some admixtures. In contrast, Ouaja *et al*. (2021) structured 304 Tunisian durum wheat landraces into 11 genetically distinct groups. But the existence of admixture of 41 genotypes into other population groups was found to be similar to our result. In addition, a more structured (K = 7) and genetically high admixture sub-populations were identified by Arora *et al*. (2014) among 319 Indian durum wheat genotypes. Marzario *et al*. (2018) also found six population clusters with few admixtures among 136 landraces and 28 varieties of Italian durum wheat genotypes. In a recent study by Kefyalew Negisho *et al*. (2021), a set of 285 durum wheat genotypes, comprising 215 Ethiopian durum wheat landraces, 10 durum wheat cultivars, 10 advanced durum wheat lines from Ethiopia, and 50 durum wheat lines from CIMMYT were genotyped using SNP markers. The whole panel was attributed to two populations representing mainly the landraces on one hand, and released, advanced and CIMMYT lines on the other hand.

7. Conclusion

The introduction of exotic durum wheat in Ethiopia to support the production of durum wheat for the processing sectors often failed to adapt to the wide agro-ecological and climatic conditions of the country (Dejene Mengistu *et al*., 2015). Hence documenting the genetic diversity of landraces collected from different agro-ecological zones of Ethiopia is important to find adaptive genotypes with optimum quality attributes for pasta and related products and to consider the selected genotypes in the breeding program of durum wheat.

The morphological study of 420 Ethiopian durum wheat genotypes identified high genetic variation among and within the studied group of genotypes in the analysis of variance and variance components. This variation may be attributed to the cultivation of durum wheat in wide agro-ecological zones of Ethiopia and/or the rare cross-pollinating character of durum wheat in the field of cultivation. The proportion of heritable variation from the observed variation for effective selection as estimated by heritability and genetic advance analysis showed greater than 40 % value of heritability for GPC, GL, TGW, GY, DH, DM, SPS and SPL with moderate to low GAM.

The significant and negative correlation of both GPC and GL with GY, and GPC with TGW was observed. Hence, breeding for the improvement of the protein content of durum wheat will decrease yield traits. Five major clusters with significant inter-cluster distance were identified using the Wards method of clustering in 420 Ethiopian durum wheat genotypes based on 11 quantitative traits. All clusters have shown unique mean value for at least one quantitative trait from which a high mean value of GL and GPC was observed in C4 and high TGW and GY belonged to C5. The PCA revealed the distinctiveness of cultivars and some genotypes of Arsi

from the other landraces. Qualitative traits also showed considerable genetic variation among geographical origins with significant variation of Shannon diversity index for KCL, VTR, and SPD. Correspondence analysis also discriminates cultivars from other landraces in terms of phenotypic classes of qualitative traits. Based on desirable phenotypic traits the top 5% genotypes were selected. The selection criteria primarily considers genotypes with white kernel color and vitreous kernel and then high GL and GPC with moderate to high TGW and GY.

In the molecular diversity study, 11 out of 14 SSR markers were highly informative (PIC>0.5) with high gene diversity (h>0.5). Of the total variation genotypes within populations contributed 88.35 % and 11.65 % of variation occurred among populations as revealed by AMOVA. The genetic distance and differentiation analysis showed the existence of high genetic distance and differentiation of Cultivars and Arsi populations from other populations which confirmed the finding of PCA and cluster analysis of quantitative morphological traits. Both NJ method of clustering and PCoA grouped 104 durum wheat genotypes into three major clusters. The population structure analysis revealed two genetic groups with some admixtures in the landraces.

8. Recommendations

Based on the finding of this study, the following key recommendations are given for consideration:

- ✓ As the studied Ethiopian durum wheat genotypes showed high genetic variation for the traits considered, breeders should consider those genotypes in durum wheat breeding programs.
- ✓ Breeding programs for the improvement of traits with high to moderate heritability value with moderate genetic advance (GL, TGW, GY and SPL) in Ethiopian durum wheat genotypes would be effective.
- ✓ Due to the significant and negative correlation of both GPC and GL with GY, and GPC with TGW in durum wheat, it is better to identify landraces having optimal protein content along with moderate to high yield potential or develop a cultivar having improved protein content and yield.
- ✓ To develop a progeny with high GPC and GL and moderate to high yield of durum wheat, it is better to cross superior individuals from C4 and C5.
- ✓ The top-performing genotypes for the studied quality traits should be assessed for technological traits variations like semolina yield, gluten and dough strength, yellow pigment content, sedimentation volume, and so on to find the very best landrace for the production of quality pasta and related products.
- ✓ The selection of genotypes based on phenotype in a narrow number and type of environment may not be repeatable under another set of environments. Hence, Ethiopian durum wheat genotypes (particularly the top 5% genotypes) for these phenotypic traits

should be studied further across environments to investigate the consistency of their performance.

- ✓ The 11 SSR markers are recommended to be used in characterizing, gene mapping, gene tagging, and other related molecular studies of durum wheat genotypes.
- ✓ The molecular diversity study using SSR markers should be supported by phenotypic and/or passport data to link the variability with the morphological and/or ecological differences, respectively.
- ✓ By adding numbers of markers, it is better to identify marker-trait associations and mapping of quantitative trait loci (QTL) for targeted phenotypic traits.

9. References

Abhinandan, K., Skori, L., Stanic, M., Hickerson, M.N., Jamshed, M. and Samuel, M.A. (2018). Abiotic Stress Signaling in Wheat – An Inclusive Overview of Hormonal Interactions During Abiotic Stress Responses in Wheat. *Front. Plant Sci*. **9**:734.

Achtar, S., Moualla, M. Y., Kalhout, A., Röder, M.S. and MirAli, N. (2010). Assessment of genetic diversity among Syrian durum (*Triticum turgidum* ssp. *durum*) and bread wheat (*Triticum aestivum* L.) using SSR markers. *Genetika*. **46**(11):1500-1506.

Adugna Tolera (2007). Feed Resources for Producing Export Quality Meat and Livestock in Ethiopia: Examples from Selected Woredas in Oromia and SNNP Regional States. Ethiopia SPS-LMM Program, Addis Ababa, Ethiopia. 77p.

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

AOAC (2016). *Official Methods of Analysis Association of Official Analysis Chemists*. 20th Edn., AOAC, Washington DC, USA.

Ahmad, S,Q., Khan, S., Ghaffar, M. and Ahmad, F. (2011). Genetic diversity analysis for yield and other parameters in maize (*Zea mays* L.) genotypes. *AJAFS.* **3**(5):385–388.

Akcura, M. (2009). Genetic variability and interrelationship among grain yield and some quality traits in Turkish winter durum wheat landraces. *Turk J Agric For*. **33**(6):547-556.

Ali, H.I. (2012). Heritability, variability, genetic correlation and path analysis for quantitative traits in durum and bread wheat under dry farming conditions. *MJA*. **40**(4):27-39.

Amallah, L., Taghouti, M., Rhrib, K., Gaboun, F., Arahou, M., Hassikou, R., and Diria, G. (2016). Validation of simple sequence repeats associated with quality traits in durum wheat. *JCSB*. **19**(2):137-150.

Arega Gashaw, Mohammed, H. and Singh, H. (2007). The selection criterion for improved grain yields in Ethiopian durum wheat genotypes. *Afr. Crop Sci. J.* **15**:25-31.

Arora, A., Kundu, S., Dilbaghi, N., Sharma, I., and Tiwari, R. (2014). Population structure and genetic diversity among Indian wheat varieties using microsatellite (SSR) markers. *AJCS*. **8**(9):1281-1289.

Asfaw Negassa, Bekele Shiferaw, Koo, J., Sonder, K., Smale, M., Braun, H., Gbegbelegbe, S., Guo, Z., Hodson, D., Wood, S., Payne, T. and Abeyo, B. (2013). *The potential for wheat production in Africa: analysis of biophysical suitability and economic profitability*. Technical report, CIMMYT, Mexico.

Bakala, H.S., Mandahal, K.S., Sarao, L.K. and Srivastava, P. (2021). Breeding Wheat for Biotic Stress Resistance: Achievements, Challenges and Prospects, pp, 1-30. **In** *Current Trends in Wheat Research*. ItechOpen.

Bayush Tsegaye and Berg, T. (2007). Utilization of durum wheat landraces in East Shewa, central Ethiopia: Are home uses an incentive for on-farm conservation? *Agric. Human Values*. **24**(2), 219–230.

Belderok B, Mesdag J, Donner D.A. (2000). *Bread-Making Quality of Wheat: A Century of Breeding in Europe, pp. 30–31*. Kluwer Academic Publisher; Dordrecht, Netherlands.

Bhandari, H.R., Bhanu, A.N., Srivastava, K., Singh, M.N. and Shreya, H.A. (2017). Assessment of genetic diversity in crop plants-an overview. *Adv. plants agric. res.* **7**(3):00255.

Bio-Rad Laboratories, Inc. https://www.bio-rad.com/en-sg/product/image-lab-software?ID=KRE6P5E8Z

Brasesco, F., Asgedom, D., Sommacal, V., Casari G. (2019). Strategic analysis and intervention plan for wheat and wheat products in the Agro-Commodities Procurement Zone of the pilot Integrated Agro-Industrial Park in Central-Eastern Oromia, Ethiopia. Addis Ababa. FAO.

Bello, D., Kadams, A.M., Simon, S.Y. and Mashi, D.S. (2007). Studies on genetic variability in cultivated sorghum (Sorghum bicolor L. Moench) cultivars of Adamawa State Nigeria. *AEJAES*. **2**(3):297–302

Bilgin, O., Korkut, K. Z., Başer, I., Dağlioğlu, O., Öztürk, I., Kahraman, T., and Balkan, A. (2010). Variation and heritability for some semolina characteristics and grain yield and their relations in durum wheat (*Triticum durum* Desf.). *World J. Agric. Res.* **6**(3):301-308.

Bilgin, O., Korkut, K.Z., Başer, I., Dağlioğlu, O., Öztürk, I, Kahraman, T., and Balkan, A. (2011). Genetic variation and inter-relationship of some morpho-physiological traits in durum wheat (*Triticum durum* (L.) Desf.). *Pak. J. Bot*. **43**(1):253-260.

Blanco, A., Mangini, G., Giancaspro, A., Giove, S., Colasuonno, P., Simeone, R., and Gadaleta, A. (2012). Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. *Molecular Breeding*. **30**(1):79-92.

Brankovic, G., Dodig, D., Zoric, M.Z., Surlan-Momirovic, G.G., Dragicevic, V., and Duric, N. (2014). Effects of climatic factors on grain vitreousness stability and heritability in durum wheat. *Turk J Agric For*. **38**(4):429-440.

Brankovic, G., Dodig, D., Pajić, V., Kandić, V., Knežević, D., Djurić, N. and Živanović, T. (2018). Genetic parameters of *Triticum aestivum* and *Triticum durum* for technological quality properties in Serbia. *Zemdirbyste*. **105**(1):39-48.

Burton C.W. and Devane E.H. 1953. Estimating heritability in tall Festuca (Restucaarundinaceae) from donar material. *Agron. J.* **45**:1476-1481.

Chahal, G.S., and Gosal, S.S. (2002). *Principles and procedures of plant breeding: Biotechnological and conventional approaches*. Alpha Science Int'l Ltd..

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R. and Chhetri, N. (2014). A meta-analysis of crop yield under climate change and adaptation. *Nat. Clim. Chang*. **4**:287–291.

Daniel Hailegiorgis, Lee, C.A. and Yun, S.J. (2017). Allelic Composition and Associated Quality Traits of the Glu-1 and Glu-3 Loci in Selected Modern Ethiopian Durum Wheat Varieties. *J. Crop Sci. Biotech.* **20**(5):387-392.

Dawit Tsegaye, Dessalegn, T., Dessalegn, Y., and Share, G. (2012). Analysis of genetic diversity in some durum wheat (*Triticum durum* Desf) genotypes grown in Ethiopia. *Afr. j. biotechnol*. **11**(40):9606-9611.

Dejene Mengistu and Pè, M.E. (2016). Revisiting the ignored Ethiopian durum wheat (*Triticum turgidum* var. *durum*) landraces for genetic diversity exploitation in future wheat breeding programs. *J. Plant Breed. Crop Sci.* **8**(4):45-59.

Dejene Mengistu, Kidane, Y.G., Fadda, C., and Pè, M.E. (2016). Genetic diversity in Ethiopian durum wheat (*Triticum turgidum* var *durum*) inferred from phenotypic variations. *Plant Genet. Res*. **16**(1):39-49.

Dejene Mengistu, Kiros, A.Y., and Pè, M.E. (2015). Phenotypic diversity in Ethiopian durum wheat (*Triticum turgidum* var. *durum*) landraces. *Crop J*. **3**(3):190-199.

Dexter, J., Williams, P., Edwards, N., and Martin, D. (1988). The relationships between durum wheat vitreousness, kernel hardness and processing quality. *J. Cereal Sci*. **7**(2):169-181.

Efrem Bechere, Belay, G., Mitiku, D., and Merker, A. (1996a). Phenotypic diversity of tetraploid wheat landraces from northern and north-central regions of Ethiopia. *Hereditas*. **124**(2):165-172.

Efrem Bechere, Hirut Kebede and Getachew Belay (2000). Durum wheat in Ethiopia. An old crop in an ancient land. IBCR, Addis Ababa, Ethiopia.

Efrem Bechere, Tesfaye Tesemma, T., and Mitiku, D. (1996b). Performance of durum wheat genotypes under two naturally waterlogged highland vertisols of Ethiopia. *EJAS*.

Ellis, R.P., Forster, B.P., Robinson, D., Handley, L.L., Gordon, D.C., Russell, J.R. and Powell, W. (2000). Wild barley: a source of genes for crop improvement in the 21st century?. *JXB*. **51**(342):9-17.

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol*. **14**(8):2611–20.

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

Ethiopian Institute of Agricultural Research (EIAR). http://www.eair.gov.et/index.php/en/research/research-centers (accessed on12 July 2020).

Endashaw Bekele (1984). Analysis of regional patterns of phenotypic diversity in the Ethiopian tetraploid and hexaploid wheats. *Hereditas*. **100**(1):131-154.

Excoffier, L., and Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res*. **10**(3):564-567.

Falconer, D.S. and Mackay, T.F.C. (1996). *Introduction to Quantitative Genetics*. 4th edn. Longman, Group Ltd., London.

FAO. Policy Support and Governance Gateway. Food Price Index (FPI). Available online: http://www.fao.org/worldfoodsituation/foodpricesindex/en/ (accessed on 12 July 2021).

Faris Hailu, Johansson E, Merker A (2010). Patterns of phenotypic diversity for phenologic and qualitative traits in Ethiopian tetraploid wheat germplasm. *Genet. Resour. Crop Evol.* **57**:781-790.

Faris Hailu (2011). Genetic Diversity and Grain Protein Composition of Tetraploid Wheat (*Triticum durum* Desf.) Germplasm from Ethiopia. Doctoral Thesis, Swedish University of Agricultural Sciences, Sweden.

Faris Hailu, Merker, A., Harjit-Singh, Getachew Belay and Johansson, E. (2006). Multivariate analysis of diversity of tetraploid wheat germplasm from Ethiopia. *Genet. Resour. Crop Evol*. **53**(6):1089-1098.

Faris Hailu, Merker, A., Getachew Belay and Johansson, E. (2005). Molecular diversity and phylogenic relationships of tetraploid wheat species as revealed by inter-simple sequence repeats (ISSR) from Ethiopia. *J. Genet. Breed*. **59**:329-338.

Farooq, S., and Azam, F. (2002). Molecular markers in plant breeding-I: Concepts and characterization. *PJBS.* **5**(10):1135-1140.

Fassil Kebebew, Yemane Tsehaye and Mcneilly, T. (2001). Diversity of durum wheat (*Triticum durum* Desf.) at in situ conservation sites in North Shewa and Bale, Ethiopia. *JAS.* **136**(4):383-392.

Feillet, P. (1988). Protein and enzyme composition of durum wheat. **In**: *Durum Chemistry and Technology*, pp. 93-119, (Fabriani, G. and Lintas, C.,eds). AACC, St. Paul, Minnesota.

Feldman, M. and Kislevb, M.E. (2007). Domestication of emmer wheat and evolution of free-threshing tetraploid wheat. *Isr. J. Plant Sci.* **55**:207–221.

Frankel, O.H., Brown, A.H., and Burdon, J.J. (1995). *The conservation of plant biodiversity*. Cambridge University Press.

Fiore, M.C., Mercati, F., Spina, A., Blangiforti, S., Venora, G., Dell'Acqua, M. and Sunseri, F. (2019). High-throughput genotype, morphology, and quality traits evaluation for the assessment of genetic diversity of wheat landraces from sicily. *Plants*. **8**(5):116.

Firdissa Eticha, Endashaw Bekele, Getachew Belay, and Borner, A. (2005). Phenotypic diversity in tetraploid wheats collected from Bale and Wello regions of Ethiopia. *Plant Genet. Res.* **3**(1), 35-43.

Fu, B., Wang, K., Dupuis, B., Taylor, D., and Nam, S. (2017). Kernel vitreousness and protein content: Relationship, interaction and synergistic effects on durum wheat quality. *J. Cereal Sci.* **78**:2-9.

Gepts, P. (2006). Plant genetic resources conservation and utilization: the accomplishments and future of a societal insurance. *Policy Crop Sci* **46**:2278–2292.

Getachew Belay, Tesfaye Tesemma, Becker H.C. and Merker, A. (1993). Variation and interrelationships of agronomic traits in Ethiopian tetraploid wheat landraces. *Euphytica.* **71**:181-188.

Gomez, K.A. and Gomez, A. (1984). *Statistical Procedures for Agricultural Research*. John Wiley and Sons. New York.

Grahmann, K., Verhulst, N., Peña, R. J., Buerkert, A., Vargas-Rojas, L., and Govaerts, B. (2014). Durum wheat (*Triticum durum* L.) quality and yield as affected by tillage–straw management and nitrogen fertilization practice under furrow-irrigated conditions. *Field Crops Res.* **164**:166-177.

Gulmezoglu, N., Alpu, O., and Ozer, E. (2010). Comparative performance of triticale and wheat grains by using path analysis. *BJAS*. **16**(4):443-453.

Habtamu Ayalew, Tsige Genet, Tadesse Dessalegn and Landuber Wondale (2011). Multivariate diversity, heritability and genetic advance in tef landraces in Ethiopia. *Afri.Crop Sci. J.* **3**(19): 201-212.

Hakan, O., Willcox, G., Graner, A., Salamini, F. and Kilian, B. (2010). Geographic distribution and domestication of wild emmer wheat (*Triticum dicoccoides*). *Genet. Resour. Crop Evol*. **58**:11–53.

Hailu Gebre-Maryam (1991). Wheat production and research in Ethiopia. **In**: *Wheat Research in Ethiopia: A Historical Perspective,* pp. 1–15, (Hailu Gebre-Mariam, Tanner, D.G. and Mengistu Huluka, eds). IAR/ CIMMYT, Addis Ababa, Ethiopia.

Hancock, J. F. (2004). *Plant Evolution and the Origin of Crop Species*. 2nd edn. CABI Publishing., Cambridge, UK.

Hruskova, M., and Švec, I. (2009). Wheat hardness in relation to other quality factors. *CJFS*. **27**(4):240-248.

Huang, C., Verrillo, F., Renzone, G., Arena, S., Rocco, M., Scaloni, A., and Marra, M. (2011). Response to biotic and oxidative stress in Arabidopsis thaliana: Analysis of variably phosphorylated proteins. *J. Proteomics*. **74**(10):1934–1949.

Husson, F., Josse, J., Le, S., Mazet, J. and Husson, M. F. (2016). Package 'FactoMineR'. *An R package*. **96**:698.

International Board for Plant Genetic Resources (IBPGR) (1985). Revised Descriptor List for Wheat (Triticum spp.). (Rome, Italy).

International Market Analysis Research Consulting (IMARC). https://www.imarcgroup.com/pasta-market (accessed in 12 July 2021).

International Grain Council (IGC). Grain market report. https://www.igc.int/en/default.aspx (accessed in 09 July 2021)

International Pasta Organisation (IPO). Annual Report. Available online: https://internationalpasta.org/annual-report/ (accessed on 12 July 2021).

Jemanesh Kifetew Haile, Hammer, K., Ayele Badebo, Nachit, M.M. and Roder, M.S. (2013). Genetic diversity assessment of Ethiopian tetraploid wheat landraces and improved durum wheat varieties using microsatellites and markers linked with stem rust resistance. *Genet. Resour. Crop Evol.* **60**:513–527.

Johnson, H.W., Robinson, H.F. and Comstock, R.F. (1955). Genotypic and phenotypic correlation in soybean and their implication in selection. *Agronomy J.* **47**(10):477-483.

Johnson, M., Kumar, A., Oladzad-Abbasabadi, A., Salsman, E., Aoun, M., Manthey, F.A., and Elias, E.M. (2019). Association mapping for 24 traits related to protein content, gluten strength, color, cooking, and milling quality using balanced and unbalanced data in durum wheat [*Triticum turgidum* L. var. *durum* (Desf).]. *Frontiers in genetics*. **10**:717.

Kabbaj, H., Sall, A.T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., Belkadi, B., Ortiz, R. and Bassi, F.M. (2017). Genetic Diversity within a Global Panel of Durum Wheat (*Triticum durum*) Landraces and Modern Germplasm Reveals the History of Alleles Exchange. *Front. Plant Sci*. 8:1277.

Kadkol G.P., and Sissons M. (2016). Durum Wheat: Overview. **In**: *Encyclopedia of Food Grains*, 2nd Edition, pp. 117-124 (Wrigley, C., Corke, H., and Seetharaman, K., Faubion, J., eds.). Oxford Academic Press, London, England.

Kassambara, A. and Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7.

Kefyalew Negisho, Shibru, S., Pillen, K., Ordon, F., and Wehner, G. (2021). Genetic diversity of Ethiopian durum wheat landraces. *Plos one*. **16**(2):e0247016.

Kehel, Z., Garcia-Ferrer, A., and Nachit, M.M. (2013). Using Bayesian and Eigen approaches to study spatial genetic structure of Moroccan and Syrian durum wheat landraces. *AJMB.* **3**:17-31.

Khan, M.A., Mohammad, F., Malik, T., Khan, A., and Abbas, S. J. (2013). Genetic divergence in f4: 6wheat lines for yield and its contributing traits. *JPBG*. **1**(3):169-175.

Khlestkina, E.K., Huang, X,Q., Quenum, F.J.B., Chebotar, S., Roder, M.S. and Börner, A. (2004) Genetic diversity in cultivated plants loss or stability? *Theor. Appl. Genet.* **108**:1466–1472.

Kirouani, A., Henkrar, F., Udupa, S. M., Boukhalfoun, L., and Bouzerzour, H. (2018). Genetic diversity in Algerian durum wheat varieties (*Triticum turgidum* L var. *durum*) using microsatellite markers. *Bioscience*. **34**(6):1575-83.

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

Kumar, P., Gupta, V.K., Misra, A.K., Modi, D.R., and Pandey, B.K. (2009). Potential of Molecular Markers in Plant Biotechnology. *Plant Omics.* **2**(4)**:** 141-162.

Kumar, S., Kumar, R. and Shamim, M.Z. (2015). Microsatellite Marker Based Characterization and Diversity Analysis of Wheat. *The Bioscan*. **10**(4): 2099-2105.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *MBE*. **35**:1547-1549.

Kumari, M., Kumar, M., Singh, V., Kumar, S.V. and Chaudhary L. (2019). Genic Microsatellite Markers for Genetic Diversity in Wheat Genotypes. *Int. J. Curr. Microbiol. App. Sci.* ***8***(9):1220-1231.

Larik, A.S., Malik, S.I., Kakar, A.A. and Naz, M.A. (2000). Assessment of heritability and genetic advance for yield components. **In** G. hirsutum. *Scient. Khyber*. **13**:39-44.

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

Lourenco, V.M., Rodrigues, P.C., Pires, A.M., Piepho, H.P., (2017). A robust DF-REML framework for variance components estimation in genetic studies. *Bioinformatics.* **33**:3584–3594.

Luo, M.C.; Yang, Z.L.; You, F.M. (2007). The structure of wild and domesticated emmer wheat populations, gene flow between them and the site of emmer domestication. *Theor. Appl. Genet*. **114**:947–959.

Maccaferri, M., Sanguineti, M. C., Demontis, A., El-Ahmed, A., Garcia del Moral, L., Maalouf, F., and Tuberosa, R. (2011). Association mapping in durum wheat grown across a broad range of water regimes. *JXB*. **62**(2):409-438.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. and Hornik, K.(2019). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0.

Mangini, G., Taranto, F., Giove, S., Gadaleta, A., and Blanco, A. (2010). Identification of durum wheat cultivars by a minimum number of microsatellite markers. *Cereal Res. Commun*. **38**(2):155-162.

Mariani, B. M., D'egidio, M.G. and Novaro, P. (1995). Durum Wheat Quality Evaluation: Influence of Genotype and Environment. *Cereal Chemistry*. **72**(2):194-197.

Marzario, S., Logozzo, G., David, J.L., Zeuli, P.S., and Gioia, T. (2018). Molecular genotyping (SSR) and agronomic phenotyping for utilization of durum wheat (*Triticum durum* Desf.) ex situ collection from Southern Italy: a combined approach including pedigreed varieties. *Genes*. **9**(10): 465.

Masood, M.A., Qamar, M. and Raza, I. (2018). Comparative efficiency of alpha lattice design versus randomized complete block design in wheat field trials. *IJSER*. **9**(11):646-650.

Medini, M., Hamza, S., Rebai, A. and Baum, M. (2005). Analysis of genetic diversity in Tunisian durum wheat cultivars and related wild species by SSR and AFLP markers. *Genet. Resour. Crop Evol.* **52**: 21–31.

Meseret Asmamaw, Gemchu Keneni and Kassahun Tesfaye (2019). Genetic Diversity of Ethiopian Durum Wheat (*Triticum durum* Desf.) Landrace Collections as Revealed by SSR Markers. *ACST*. **7**:1

Meseret Asmamaw Wondifaw, Gemechu Keneni and Kassahun Tesfaye (2020). Genetic diversity of Ethiopian durum wheat (*Triticum durum* Desf) landrace collections as revealed by morphological markers. *JPBCS.* **12**(4):258-268.

Mickky, B.M. and Aldesuquy, H.S. (2017). Impact of osmotic stress on seedling growth observations, membrane characteristics and antioxidant defense system of different wheat genotypes. *Egypt. J. Basic Appl. Sci*. **4**:47–54.

Milatovic, D., Nikolic, D. and Durovic, D. (2010). Variability, heritability and correlation of some factors affecting productivity in peach. *Horticultural Science*. **27**(3):79-87.

Mohammed Abinsa**,** Geremew Bultosa and Amsalu Ayana (2012). Variation and Associations of Quality Parameters in Ethiopian Durum Wheat (*Triticum turgidum* L. var. *durum*) Genotypes. *IJPBG.* **6**(1):17-31.

Mondini, L., Farina, A., Porceddu, E., and Pagnotta, M. A. (2010). Analysis of durum wheat germplasm adapted to different climatic conditions. *Ann. Appl. Biol.* **156**(2):211-219.

Moradi, N., Badakhshan, H., Mohammadzadeh, H., Zakeri, M., and Mirzaghaderi, G. (2014). Assessment of genetic diversity and identification of SSR markers associated with grain iron content in Iranian prevalent wheat genotypes. *J. Plant Mol. Breed.* **2**(1):64-73.

Morris, R. and Sears, E.R. (1967). The cytogenetics of wheat and its relatives. **In**: *Wheat and Wheat Improvement*, pp. 19-87, (Quisenberry, K.S. and Reitz, L.P., eds.). ASA Madison.

Motalebi, M., Keshavarzi, M. and Naghavi, M.R. (2007). Glitenin Subunit composition in Durum (*Triticum durum*) landraces and cultivars. *Asian J. Plant Sci.* **6**(2):399-402.

Mulatu Aberra Ebsa and Tilahun Bayisa Worku (2017). Registration of "Bullaallaa" Newly Released Durum Wheat (*Triticum turgidum* L.) Variety for Bale Mid and Highland Areas. *AEJAES.* **17**(4): 349-353.

Mulugeta Negassa (1986). Patterns of diversity of Ethiopian wheats (*Triticum* spp.) and a gene center for quality breeding. *Plant Breed*. **97**:147-162.

Mun, J.H., Kim, D.J., Choi, H.K., Gish, J., Debelle, F., Mudge, J., Denny, R., Endre, G., Saurat, O., Dudez, A. M., Kiss, G. B., Roe, B., Young, N. D. and Cook, D. R. (2006). Distribution of microsatellites in the genome of Medicago truncatula: a resource of genetic markers that integrate genetic and physical maps. *Genetics. Apr*. **172**(4):2541-55.

Nadeem, M.A., Nawaz, M.A., Shahid, M.Q., Doğan, Y., Comertpay, G., Yıldız, M. and Baloch, F. S. (2018). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnol. Biotechnol. Equip.* **32**(2): 261-285.

Najeeb S., Rather, A.G., Parray, G.A., Sheikh F.A. and Razvi, S.M. (2009). Studies on genetic variability, genotypic correlation and path coefficient analysis in maize under high altitude temperate ecology of Kashmir. *MNL*. **83**:1-8.

Neethirajan, S., Karunakaran, C., Symonsc, C. and Jayas, D.S. (2006). Classification of vitreousness in durum wheat using soft X-rays and transmitted light images. *Comput. Electron. Agric*. **53**:71–78.

Oliveira, H.R., Campana, M.G., Jones, H., Hunt, H.V., Leigh, F., Redhouse, D.I., and Jones, M.K. (2012). Tetraploid wheat landraces in the Mediterranean basin: taxonomy, evolution and genetic diversity. *Plos one*. ***7***(5):e37063.

Ouaja, M., Bahri, B. A., Aouini, L., Ferjaoui, S., Medini, M., Marcel, T. C., and Hamza, S. (2021). Morphological characterization and genetic diversity analysis of Tunisian durum wheat (*Triticum turgidum* var. *durum*) genotypes. *BMC Genomic Data*. **22**(1):1-17.

Oumer, O.A., Dagne, K., Feyissa, T., Tesfaye, K., Durai, J., and Hyder, M. Z. (2020). Genetic diversity, population structure, and gene flow analysis of lowland bamboo

[Oxytenanthera abyssinica (A. Rich.) Munro] in Ethiopia. *Ecol. Evol*. **10**(20):11217-11236.

Padalino, L., Del Nobile, M.A., la Gatta, B., Rutigliano, M., Di Luccia, A., and Conte, A. (2019). Effects of microwave treatment of durum wheat kernels on quality characteristics of flour and pasta. *Food chemistry*. **283**:454-461.

Pasha, I., Anjum, F.M., and Morris, C.F. (2010). Grain hardness: a major determinant of wheat quality. *FSTI*. **16**(6):511-522.

Peakall, R.and Smouse, R., (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. **28**(19):2537-2539.

Pecetti, L., Annicchiarico, P. and Damania, AB. (1992). Biodiversity in a germplasm collection of durum wheat. *Euphytica*. **60**:229- 238.

Peterson, B.G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., Zivot, E. and Wuertz, D. (2018). Package 'performanceanalytics'. *R Team Cooperation*. **3**:13-14.

Pitz, W. (1992). Durum wheat/semolina/farina/pasta quality. North Dakota State University.

Pritchard, J. K., Wen, W. and Falush, D. (2010). Documentation for STRUCTURE software: Version 2. University of Chicago, Chicago, *IL*.

Prasad, M., Kumar, N., Kulwal, P., Röder, M., Balyan, H., Dhaliwal, H., and Gupta, P. (2003). QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. *Theor. Appl. Genet*. **106**(4):659-667.

Quinn, G.P. and Keough M (2002). *Experimental Design and Data Analysis for Biologists Anessential*. Cambridge University Press, New York..

R Core Team (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria.

Rachon, L. and Szumilo, G. (2009). Yield of winter durum wheat (*Triticum durum* Desf.) lines in condition of different protection level of plants. *Acta. Sci. Pol. Agricultura*. **8**:15–22.

Rajendrakumar, P., Biswal, A. K., Balachandran, S. M., Srinivasarao, K., and Sundaram, R. M. (2007). Simple sequence repeats in organellar genomes of rice: frequency and distribution in genic and intergenic regions. *Bioinformatics*. **23**(1):1-4.

Ramya, P., Chaubal, A., Kulkarni, K., Gupta, L., Kadoo, N., Dhaliwal, H. S., and Gupt, V. (2010). QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (*Triticum aestivum* L.). *JAG*. **51**(4):421-429.

Rapp, M., Sieber, A., Kazman, E., Leiser, W., Würschum, T., and Longin, C. (2019). Evaluation of the genetic architecture and the potential of genomics-assisted breeding of quality traits in two large panels of durum wheat. *Theor. Appl. Genet*. **132**(6):1873-1886.

Rencher, A.C. (2002). *Methods of Multivariate Analysis*, 2nd edn. Brigham Young University, John Wiley and Sons, Inc, London.

Rizwan, S., Ahmad, I., Ashraf, M., Iqbal-Mirza, J., Mustafa-Sahi, G., Atiq-ur-Rahman, R. and Mujeeb-Kazi, A. (2007). Evaluation of Synthetic Hexaploid Wheats (*Triticum turgidum* L. x *Aegilops tauschii* L.) and their Durum Parents for Stripe Rust (*Puccinia striiformis* Westend. f. sp. *tritici* Erikson) Resistance. *Rev. mex. fitopatol*. **25**(2):152-160.

Rocco, M., Tartaglia, M., Izzo, F. P., Varricchio, E., Arena, S., Scaloni, A., and Marra, M. (2019). Comparative proteomic analysis of durum wheat shoots from modern and ancient cultivars. *Plant Physiol. Biochem*. **135**:253-262.

Roder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, P. and Ganal, M.W. (1998). A microsatellite map of wheat. *Genetics.* **149**:2007- 2023.

Rohlf, F.J. (2002). NTSYS-pc: numerical taxonomy system ver. 2.1. *Setauket, NY: Exeter Publishing Ltd.*

Sacchetti, G., Cocco, G., Cocco, D., Neri, L. and Mastrocola, D. (2011). Effect of semolina particle size on the cooking kinetics and quality of spaghetti. *Procedia Food Sci*. **1**:1740-1745.

Sall, T.A., Chiari, T., Legesse, W., Seid-Ahmed, K., Ortiz, R., Van Ginkel, M., and Bassi, F.M. (2019). Durum wheat (*Triticum durum* Desf.): Origin, cultivation and potential expansion in Sub-Saharan Africa. *Agronomy*. **9**(5):263.

Sayaslan, A., Koyuncu, M., Yildirim, A., Eserkaya Gulec, T., Ates Sonmezoglu, O. and Kandemir, N. (2012). Some quality characteristics of selected durum wheat (*Triticum durum*) landraces. *Turk. J. Agric. For*. **36**:749-756.

Scrucca L., Fop M., Murphy T.B. and Raftery A.E. (2016) mclust5: clustering, classification and density estimation using Gaussian finite mixture models *The R Journal*. **8**(1):289-317.

Seifu Tsegaye, Tesfaye Tesemma and Getachew Belay (1996). Relationships among tetraploid wheat *(Triticum turgidum* L.) landrace populations revealed by isozyme markers and agronomic traits. *Theor. Appl. Genet.* **93**:600-605.

Singh, R. P., Huerta-Espino, J., Pfeiffer, W. and Figueroa-Lopez, P. (2004). Occurrence and impact of a new leaf rust race on durum wheat in northwestern Mexico from 2001 to 2003. *Plant Dis*. ***88***(7):703-708.

Sintayehu Alamerew, Chebotar, S., Huang, X., Röder, M., and Börner, A. (2004). Genetic diversity in Ethiopian hexaploid and tetraploid wheat germplasm assessed by microsatellite markers. *Genet. Resour. Crop Evol*. **51**(5):559-567.

Singh, K.B. (1990). Prospects of developing new genetic material and breeding methodologies for chickpea improvement. *CIHEAM–Options Mediterraneennes.* **9**:43-50.

Singh, B.D. (2007). *Plant Breeding: Principles and Methods*, 7th Edn. Kalyani Publishers, New Delhi, India.

Singh, R.K. and Chaudhary (1977). *Biometrical Methods in Quantitative Genetic Analysis*. Kalyani publishers, New Delhi-Ludhiana India.

Sissons, M.J. (2004). Pasta. **In**: *Encyclopaedia of Grain Science*, pp. 409-418, (Wrigley, C., Corke, H., Walker, C., eds). Tamworth Agricultural Institute, Tamworth, Australia.

Sissons M. (2016) Pasta. **In**: *Encyclopedia of Food Grains*, pp. 79-89, (Wrigley, C., Corke, H., and Seetharaman, K., Faubion, J., eds.). Tamworth Agricultural Institute, Tamworth, Australia.

Slim, A., Piarulli, L., Chennaoui Kourda, H., Rouaissi, M., Robbana, C., Chaabane, R., and Mangini, G. (2019). Genetic structure analysis of a collection of Tunisian durum wheat Germplasm. *IJMS*. **20**(13):3362.

Sourour, A. and Hajer, S.A. (2008). Distribution and phenotypic variability aspects of some quantitiative traits among durum wheat accessions. *Afr. Crop Sci. J.* **16**(4):219-224.

Sourour, A., Othmani, A., Bechrif, S., Rezgui, M., and Ben Younes, M. (2018). Correlation between agronomical and quality traits in durum wheat (*Triticum durum* Desf.) germplasm in semi-arid environment. *Adv. plants agric. res*. **8**:612-615.

Taneva, K., Bozhanova,V. and Dragov, V.B. (2019a). Assessment of Diversity and Association between Agronomic and Quality Traits in an assortiment of Durum Wheat Genotypes. *IJIAAR*. **3**(3):480-490.

Taneva, K., Bozhanova, V. and Petrova, I. (2019b). Variability, heritability and genetic advance of some grain quality traits and grain yield in durum wheat genotypes. *BJAS.* **25**(2):288–295.

Terletskaya, N.V., Shcherban, A.B., Nesterov, M.A., Perfil'ev, R.N., Salina, E.A., Altayeva, N.A. and Blavachinskaya, I.V. (2020). Drought Stress Tolerance and Photosynthetic Activity of Alloplasmic Lines *T. dicoccum* x *T. aestivum*. *Int. J. Mol. Sci*. **21**:3356.

Tesfaye Tesemma, Getachew Belay, and Melaku Worede (1991). Morphological diversity in tetraploid wheat landrace populations from the central highlands of Ethiopia. *Hereditas*. **114**(2):171-176.

Tesfaye Wolde, Firdisa Eticha, Sentayehu Alamerew, Ermias Assefa and Dargicho Dutamo (2016a). Genetic variability, heritability and genetic advance for yield and yield related traits in Durum wheat (*Triticum durum* L.) genotypes. *SJAR*. **5**(3):042-047.

Tesfaye Wolde, Firdisa Eticha, Sentayehu Alamerew, Ermias Assefa, Dargicho Dutamo and Birhanu Mecha (2016b). Trait associations in some durum wheat (*Triticum durum* L.) genotypes among yield and yield related traits at Kulumsa, south eastern Ethiopia. *Adv Crop Sci Tech.* **4**:234.

Tester, M. and Langridge, P. (2010). Breeding Technologies to Increase Crop Production in a Changing World. *Science*. **327**:818-822.

Todorovska, E., Christov, N., Slavov, S., Christova, P. and Vassilev, D. (2009). Biotic stress resistance in wheat—breeding and genomic selection implications. *Biotechnol. Biotechnol. Equip*. **23**(4):1417-1426.

Tuhina-Khatun, M., Newaz, M.A. and Bari, M.A. (2007). Combining ability and heritability estimates in F2 diallel population of spring wheat under interacting environments. *BARI*. **34**:75–82.

Turnbull, K. (2001). Quality assurance in a dry pasta factory. **In**: *Pasta and Semolina Technology*, pp. 181-221, (Kill, R.C. and Turnbull, K., eds). Blackwell Scientific, Oxford, London, England.

Uddin, M.S. and Boerner, A., (2008). Genetic Diversity in Hexaploid and Tetraploid Wheat Genotypes Using Microsatellite Markers. *Plant Tissue Cult. and Biotech.* **18**(1):65-73.

USDA, (2015). Grain Inspection, Packers and Stockyards Administration. Federal Grain Inspection Service. USDA, Washington, USA.

Vaiman, D., Mercier, D., Moazami-Goudarzi, K., Eggen, A., Ciampolini, R., Lépingle, A., and Guérin, G. (1994). A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mammalian Genome*. **5**(5):288-297.

van de Wouw, M., van Hintum, T., Kik, C., van Treuren, R. and Visser, B. (2010). Genetic diversity trends in twentieth century crop cultivars:a meta-analysis. *Theor Appl. Genet.* **120**:1241–1252.

Vavilov N.I. (1992). *Origin and geography of cultivated plants*. Cambridge University Press, Cambridge.

Wang, K., and Fu, B.X. (2020). Inter-Relationships between Test Weight, Thousand Kernel Weight, Kernel Size Distribution and Their Effects on Durum Wheat Milling, Semolina Composition and Pasta Processing Quality. *Foods*. **9**(9):1308.

Weegels, P.L., Hamer, R.J., and Schofield, J. 1. (1996). Functional properties of wheat glutenin. *J. Cereal Sci.* **23**(1):1-17.

World Bank, (2018). Cereal Market Performance in Ethiopia: Policy Implications for Improving Investments in Maize and Wheat Value Chains. Agriculture Global Practice, GFA13.

Yadav, S., Singh. K., Pandey, P. and Singh, S. (2015). Genetic Variability and Direct Selection Criterion for Seed Yield in Segregating Generations of Barley (Hordeum vulgareL.). *Ann.J.Plant Sci*. **6**:1543-1549.

Yagdi, K. and Sozen, E. (2009). Heritability, variance components and correlations of yield and quality traits in durum wheat (*Triticum durum* Desf.). *Pak. J. Bot.* **41**(2):753-759.

Yifru Teklu, Hammer, K., Huang X.Q. and Roder, M.S. (2006). Analysis of microsatellite diversity in Ethiopian tetraploid wheat landraces. *Genet. Resour. Crop Evol.* **53**:1115–1126.

Yifru Teklu, and Hammer, K. (2006). Farmers' perception and genetic erosion of tetraploid wheat landraces in Ethiopia. *Genet. Resour. Crop Evol.* **43**, 395- 407.

Yohannes Alemu, Anley, A.M., and Abebe, T.D. (2020). Genetic variability and association of traits in Ethiopian durum wheat (*Triticum turgidium L.* var. *durum*) landraces at Dabat Research Station, North Gondar. *Cogent Food Agric*. **6**(1):1778604.

Zhang, W., Chao, S., Manthey, F., Chicaiza, O., Brevis, J. C., Echenique, V., and Dubcovsky, J. (2008). QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theor. Appl. Genet.* **117**(8):1361-1377.

Zilic, S., Barać, M., Pesic, M., Dodig, D., and Ignjatovic-Micic, D. (2011). Characterization of proteins from grain of different bread and durum wheat genotypes. *IJMS*. **12**(9):5878-5894.

Zohary, D. (1970). Centers of Diversity and Centers of Origin. **In**: *Genetic Resources in Plants-Their Exploration and Conservation,* 11, pp. 33-42, (Frankel, O.H. and Bennett, E., eds). Blackwell Science Publishing, Oxford, London, England.

10.Appendices

Appendix 1. Description of 420 Ethiopian durum wheat genotypes used in the phenotypic diversity study.

GEN. code	ACCN. name	Geographical origin
1	5328	Bale-Hararge
2	5309	North Shewa
3	5332	Gondar-Wello
4	5330	Bale-Hararge
5	5087	Bale-Hararge
6	5164	East Shewa
7	5001	East Shewa
8	5107	Bale-Hararge
9	5109	Bale-Hararge
10	5204	Gojam
11	5228	Bale-Hararge
12	5473	North Shewa
13	5257	North Shewa
14	5256	North Shewa
15	5259	North Shewa
16	5255	North Shewa
17	5168	North Shewa
18	5254	North Shewa
19	5014	West Shewa
20	5020	East Shewa
21	5468	Tigray
22	5471	Tigray
23	5260	North Shewa
24	5261	North Shewa
25	5338	Gojam
26	5336	SNNP
27	5056	Arsi
28	5291	Gojam
29	5343	Gondar-Wello
30	5601	SNNP
31	5267	East Shewa
32	5163	Arsi
33	5246	East Shewa
34	5192	SNNP
35	5071	Arsi
36	5245	East Shewa
37	5070	Bale-Hararge
38	5242	East Shewa
39	5023	West Shewa
40	5388	Bale-Hararge
41	5078	West Shewa
42	5348	Tigray
43	5344	East Shewa
44	5342	Gondar-Wello
45	5340	Gojam
46	5052	East Shewa
47	5049	East Shewa
48	5051	North Shewa
49	5207	Gojam
50	5623	East Shewa
51	5558	SNNP
52	5504	East Shewa
53	5188	SNNP
54	5278	Bale-Hararge
55	5281	East Shewa
56	5627	North Shewa
57	5554	West Shewa
58	5516	Tigray
59	5026	East Shewa
60	5057	Arsi
61	5104	Bale-Hararge
62	5288	West Shewa
63	5617	Gojam
64	5216	Gondar-Wello
65	5577	Tigray
66	5482	Arsi
67	5573	Tigray
68	5552	SNNP
69	5236	North Shewa
70	5467	Tigray
71	5009	West Shewa
72	5383	Bale-Hararge
73	5162	Tigray
74	5006	Bale-Hararge
75	5169	North Shewa
76	5576	North Shewa
77	5181	East Shewa
78	5239	Gojam
79	5375	East Shewa
80	5574	East Shewa
81	5561	Gondar-Wello
82	5476	Tigray
83	5215	Gondar-Wello
84	5294	Gondar-Wello
85	5286	Tigray
86	5452	North Shewa
87	5451	North Shewa
88	5364	East Shewa
89	5365	East Shewa
90	5220	Gondar-Wello
91	5519	SNNP
92	5523	SNNP
93	5202	Gojam
94	5522	East Shewa
95	5182	East Shewa
96	5174	East Shewa
97	5579	Gojam
98	5397	Gojam
99	5526	Tigray

100	5410	West Shewa
101	5414	Gojam
102	5146	Bale-Hararge
103	5155	North Shewa
104	5041	Bale-Hararge
105	5572	East Shewa
106	5453	Tigray
107	5102	Gojam
108	5420	East Shewa
109	5119	Bale-Hararge
110	5422	East Shewa
111	5583	East Shewa
112	5136	Tigray
113	5117	Bale-Hararge
114	5612	North Shewa
115	5158	North Shewa
116	5459	Gondar-Wello
117	5394	Tigray
118	5393	Gondar-Wello
119	5142	North Shewa
120	5327	Bale-Hararge
121	5172	North Shewa
122	5610	North Shewa
123	5602	SNNP
124	5305	East Shewa
125	5371	North Shewa
126	5326	West Shewa
127	5544	SNNP
128	5299	Gondar-Wello
129	5219	Gondar-Wello
130	5098	Bale-Hararge
131	5306	East Shewa
132	5159	North Shewa
133	5183	East Shewa
134	5184	East Shewa
135	5048	East Shewa
136	5140	North Shewa
137	7077	Arsi
138	5465	East Shewa
139	5076	North Shewa
140	5017	Bale-Hararge
141	5005	Bale-Hararge
142	5062	East Shewa
143	5591	East Shewa
144	5073	Gondar-Wello
145	5191	SNNP
146	5461	SNNP
147	5551	Tigray
148	5597	North Shewa
149	5518	SNNP
150	5047	North Shewa
151	5213	Gondar-Wello
152	5241	East Shewa
153	5165	West Shewa
154	5011	Bale-Hararge
155	5039	East Shewa
156	5433	East Shewa
157	5571	East Shewa
158	5401	East Shewa
159	5177	East Shewa
160	5230	West Shewa
161	5289	West Shewa
162	5426	SNNP
163	5272	Arsi
164	5095	Bale-Hararge
165	5372	East Shewa
166	5015	North Shewa
167	5434	East Shewa
168	5249	East Shewa
169	5189	SNNP
170	5066	Bale-Hararge
171	5190	SNNP
172	5025	West Shewa
173	5562	Gondar-Wello
174	5251	North Shewa
175	5030	East Shewa
176	5563	East Shewa
177	5457	East Shewa
178	5506	East Shewa
179	5018	Bale-Hararge
180	5035	East Shewa
181	5234	West Shewa
182	5125	Gondar-Wello
183	5543	SNNP
184	5065	Bale-Hararge
185	5594	Tigray
186	5619	Bale-Hararge
187	5069	Bale-Hararge
188	5387	Bale-Hararge
189	5100	Bale-Hararge
190	5565	East Shewa
191	5567	East Shewa
192	5432	East Shewa
193	2211	Tigray
194	4649	Tigray
195	5020	East Shewa
196	5043	West Shewa
197	5044	East Shewa
198	5057	Arsi
199	5140	North Shewa
200	5141	North Shewa
201	5142	North Shewa
202	5143	North Shewa
203	5149	West Shewa
204	5152	SNNP
205	5158	North Shewa
206	5168	North Shewa
207	5169	North Shewa
208	5171	North Shewa
209	3540	Tigray
210	5179	East Shewa
211	5181	East Shewa
212	5182	East Shewa
213	5183	East Shewa

214	5197	Gojam
215	5198	Gojam
216	5214	Gondar-Wello
217	5342	Gondar-Wello
218	5344	East Shewa
219	5354	West Shewa
220	5369	North Shewa
221	5373	East Shewa
222	5434	East Shewa
223	5441	East Shewa
224	5465	East Shewa
225	5470	Gondar-Wello
226	5472	North Shewa
227	5491	East Shewa
228	5492	East Shewa
229	5502	East Shewa
230	5504	East Shewa
231	5507	East Shewa
232	5510	North Shewa
233	5515	East Shewa
234	5526	Tigray
235	5537	West Shewa
236	5572	East Shewa
237	5576	North Shewa
238	7019	Arsi
239	5581	North Shewa
240	5582	East Shewa
241	5586	North Shewa
242	5591	East Shewa
243	5593	North Shewa
244	5597	North Shewa
245	5600	North Shewa
246	5609	North Shewa
247	5618	North Shewa
248	5627	North Shewa
249	5642	Gondar-Wello
250	5653	Tigray
251	5666	East Shewa
252	5669	Tigray
253	5707	North Shewa
254	5729	East Shewa
255	5892	North Shewa
256	5893	North Shewa
257	5898	SNNP
258	5909	North Shewa
259	5485	Gondar-Wello
260	5103	Bale-Hararge
261	5923	Arsi
262	6102	West Shewa
263	6914	Gojam
264	6933	Bale-Hararge
265	6936	Bale-Hararge
266	6955	Arsi
267	6968	Gojam
268	6971	Gojam
269	6974	Gojam
270	6975	Gojam
271	6983	Arsi
272	6987	Arsi
273	6988	Arsi
274	5568	East Shewa
275	7000	Arsi
276	7002	Arsi
277	7003	Arsi
278	7004	Arsi
279	7007	Arsi
280	7009	Arsi
281	7010	Arsi
282	7014	Arsi
283	7015	Arsi
284	7018	Arsi
285	7020	Arsi
286	7031	Arsi
287	5548	East Shewa
288	7046	Arsi
289	7050	Arsi
290	7056	Arsi
291	7060	Arsi
292	7063	Arsi
293	7064	Arsi
294	7069	Arsi
295	7075	Arsi
296	7076	Arsi
297	7078	Arsi
298	7082	Bale-Hararge
299	7083	Bale-Hararge
300	7084	Arsi
301	7104	North Shewa
302	7133	North Shewa
303	7135	North Shewa
304	7150	Arsi
305	7201	East Shewa
306	7205	East Shewa
307	7207	West Shewa
308	7209	West Shewa
309	7210	West Shewa
310	7218	West Shewa
311	7242	Arsi
312	7295	Gondar-Wello
313	7313	Arsi
314	7317	Arsi
315	7343	Bale-Hararge
316	7375	Gondar-Wello
317	7378	Gondar-Wello
318	7464	Gondar-Wello
319	7477	Gondar-Wello
320	5269	East Shewa
321	7532	Gondar-Wello
322	7568	Gondar-Wello
323	7569	Gondar-Wello
324	7572	Gondar-

		Wello
325	7576	Gondar-Wello
326	7578	Gondar-Wello
327	7580	Gondar-Wello
328	7581	Gondar-Wello
329	7626	North Shewa
330	7629	North Shewa
331	7631	North Shewa
332	7641	Gojam
333	7647	Gojam
334	7649	Gojam
335	7664	Gojam
336	7666	Gojam
337	7673	Gojam
338	7683	Gojam
339	7710	Bale-Hararge
340	7712	North Shewa
341	7713	North Shewa
342	7798	Gojam
343	7801	Gojam
344	7822	Gojam
345	7823	Gojam
346	3540	Tigray
347	7826	Gojam
348	7827	Gojam
349	7828	Gojam
350	7832	Gojam
351	7880	East Shewa
352	7999	West Shewa
353	8072	Bale-Hararge
354	214370	East Shewa
355	222393	Arsi
356	222427	Arsi
357	226225	East Shewa
358	226241	North Shewa
359	7832	Gojam
360	226897	West Shewa
361	230678	Bale-Hararge
362	235051	Gondar-Wello
363	238891	Bale-Hararge
364	239693	Bale-Hararge
365	239694	Bale-Hararge
366	239711	Bale-Hararge
367	242779	East Shewa
368	242781	East Shewa
369	242782	East Shewa
370	242783	East Shewa
371	242784	East Shewa
372	242785	East Shewa
373	242786	East Shewa
374	242787	East Shewa
375	242789	East Shewa
376	242790	East Shewa
377	242791	East Shewa
378	242792	East Shewa
379	242793	East Shewa
380	243698	Gondar-Wello
381	243700	Gondar-Wello
382	243701	Gondar-Wello
383	243703	Gondar-Wello
384	243706	Gondar-Wello
385	243717	Tigray
386	274497	Arsi
387	Yerer	Cultivar
388	Kilinto	Cultivar
389	Quamy	Cultivar
390	Mangudo	Cultivar
391	Mukiye	Cultivar
392	Ginchi	Cultivar
393	Dembi	Cultivar
394	Bichena	Cultivar
395	Robe	Cultivar
396	Don-matteo	Cultivar
397	Alem-tena	Cultivar
398	Ld-357	Cultivar
399	Tesfaye	Cultivar
400	Werer	Cultivar
401	Hitosa	Cultivar
402	Utuba	Cultivar
403	Gerado	Cultivar
404	Assasa	Cultivar
405	Arandato	Cultivar
406	Tob-66	Cultivar
407	Boohai	Cultivar
408	Cocorit-71	Cultivar
409	Foka	Cultivar
410	Ude	Cultivar
411	Bakalcha	Cultivar
412	Ejersa	Cultivar
413	Tate	Cultivar
414	Oda	Cultivar
415	Toltu	Cultivar
416	Ebsa	Cultivar
417	Dire	Cultivar
418	Ilani	Cultivar
419	Leliso	Cultivar
420	Bulala	Cultivar

GEN is genotype and ACCN is genotype.

Note: the source of all landraces is EBI and the cultivars were collected from DzARC (387-410) and SARC (411-420).

Appendix 2. Description of 104 Ethiopian durum wheat genotypes used in the molecular diversity study.

GEN Code	ACCN Name	Population
1	5597	Semen Shewa
2	5168	Semen Shewa
3	5047	Semen Shewa
4	7104	Semen Shewa
5	5309	Semen Shewa
6	5609	Semen Shewa
7	5251	Semen Shewa
8	5600	Semen Shewa
9	5892	Semen Shewa
10	5354	West Shewa
11	7210	West Shewa
12	5537	West Shewa
13	5149	West Shewa
14	5554	West Shewa
15	5043	West Shewa
16	226897	West Shewa
17	5025	West Shewa
18	5617	West Gojam
19	5198	West Gojam
20	5202	West Gojam
21	5340	West Gojam
22	5207	West Gojam
23	5397	West Gojam
24	5204	West Gojam
25	5291	West Gojam
26	6974	East Gojam
27	7801	East Gojam
28	7823	East Gojam
29	7826	East Gojam
30	7798	East Gojam
31	7822	East Gojam
32	7828	East Gojam
33	6975	East Gojam
34	5098	East Harerge
35	5278	East Harerge
36	5103	East Harerge
37	5018	East Harerge
38	5119	East Harerge
39	5095	East Harerge
40	5104	East Harerge
41	5109	East Harerge
42	5294	North Gondar
43	5213	North Gondar
44	5214	North Gondar
45	5470	North Gondar
46	5342	North Gondar
47	5219	North Gondar
48	5125	North Gondar
49	5342	North Gondar
50	8072	Bale
51	238891	Bale
52	239693	Bale
53	5006	Bale
54	5146	Bale
55	5011	Bale
56	5017	Bale
57	7031	Arsi
58	7014	Arsi
59	222393	Arsi
60	7046	Arsi
61	7317	Arsi
62	7069	Arsi
63	5163	Arsi
64	7532	South Wello
65	7568	South Wello
66	243698	South Wello
67	7378	South Wello
68	7581	South Wello
69	7580	South Wello
70	7375	South Wello
71	3540	Central Tigray
72	5453	Central Tigray
73	5669	Central Tigray
74	243717	Central Tigray
75	5653	Central Tigray
76	5468	Central Tigray
77	2211	Central Tigray
78	5348	Southern Tigray
79	5136	Southern Tigray
80	3540	Southern Tigray
81	5573	Southern Tigray
82	5286	Southern Tigray
83	5394	Southern Tigray
84	5526	Southern Tigray
85	Werer	Cultivar
86	Robe	Cultivar
87	Ejersa	Cultivar
88	Tate	Cultivar
89	Arand-ato	Cultivar
90	Yerer	Cultivar
91	Ilani	Cultivar
92	Boohai	Cultivar
93	Kilinto	Cultivar
94	Tob-66	Cultivar
95	5568	Misrak Shewa
96	5432	Misrak Shewa
97	242791	Misrak Shewa
98	5574	Misrak Shewa
99	5623	Misrak Shewa
100	214370	Misrak Shewa
101	5548	Misrak Shewa
102	5502	Misrak Shewa
103	5567	Misrak Shewa
104	5420	Misrak Shewa

GEN is genotype and ACCN is genotype.

Note: the source of all landraces is EBI and the cultivars were collected from DzARC (387-410) and SARC (411-420).

Appendix 3. Extended electrophoresed gel images of SSR markers of Xwmc24 (a) and Xgwm513 (b).