GENETIC DIVERSITY AND GENOME-WIDE ASSOCIATION MAPPING IN TETRAPLOID WHEAT (*TRITICUM TURGIDUM SPP.*) GERMPLASM OF ETHIOPIA

PHD DISSERTATION

MIHERETU FUFA GELETA

HAWASSA UNIVERSITY, HAWASSA, ETHIOPIA

OCTOBER 2024

GENETIC DIVERSITY AND GENOME-WIDE ASSOCIATION MAPPING IN TETRAPLOID WHEAT (*Triticum turgidum* spp.) GERMPLASM OF ETHIOPIA

MIHERETU FUFA GELETA

A THESIS SUBMITTED TO SCHOOL OF PLANT AND HORTICULTURE SCIENCES,
HAWASSA COLLEGE OF AGRICULTURE, SCHOOL OF GRADUATE STUDIES, HAWASSA UNIVERSITY,
HAWASSA, ETHIOPIA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN PLANT SCIENCES
(SPECIALIZATION: PLANT BIOTECHNOLOGY)

MAJOR SUPERVISOR: ANDARGACHEW GEDEBO (PhD, Asso. Prof)

CO- SUPERVISOR: TESFAYE LETTA (PhD, Lead Researcher)

OCTOBER, 2024

ADVISORS' APPROVAL SHEET

SCHOOL OF GRADUATE STUDIES HAWASSA UNIVERSITY
ADVISORS' APPROVAL SHEET
(SUBMISSION SHEET-1)

This is to certify that the PhD thesis entitled "**Genetic Diversity and Genome Wide Association Mapping in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia**" submitted in partial fulfilment of the requirements for the Degree of **Doctor of Philosophy (PhD)** with specialization in Plant Biotechnology, to graduate program of the School of Plant and Horticultural Sciences, College of Agriculture, and has been carried out by **Miheretu Fufa Geleta**, (ID. No. PhD/PBioR/0006/12), under our supervision. Therefore, we recommend that the student has fulfilled the requirements and hence hereby can submit the thesis to the school.

Andargachew Gedebo (PhD)	_____	_____
Name of major supervisor	signature	Date
Tesfaye Letta (PhD)	_____	_____
Name of co- supervisor	signature	Date

EXAMINERS' APPROVAL SHEET
SCHOOL OF GRADUATE STUDIES

HAWASSA UNIVERSITY

EXAMINERS' APPROVAL SHEET

(SUBMISSION SHEET-2)

We, the undersigned, members of the Board of Examiners of the final open defense by **Miheretu Fufa Geleta** have read and evaluated his PhD thesis entitled "**Genetic Diversity and Genome-Wide Association Mapping in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia**", and examined the candidate. This is, therefore, to certify that the thesis has been accepted in partial fulfilment of the requirements for the Doctor of philosophy degree (PhD) in specialization of plant biotechnology.

Dr. Andargachew Gedebo (PhD)	_____	_____
Name of Major supervisor	Signature	Date
Dr. Hewan Demissie (PhD)	_____	_____
Name of Internal Examiner	Signature	Date
Professor Teklehaimanot (PhD)	_____	_____
Name of External examiner	Signature	Date
Dr. Yayis Rezene (PhD)	_____	_____
Name of External examiner	Signature	Date
Dr. Hewan Demissie (PhD)	_____	_____
Name of the Chairperson	Signature	Date
_____	_____	_____
SGS Approval	Signature	Date

Final approval and acceptance of the thesis is contingent upon the submission of the final copy of the thesis to the School of Graduate Studies (SGS) through the School Graduate Committee (SGC) of the candidate's department.

DECLARATION

I hereby declare that this PhD dissertation in Plant Biotechnology is my original work and has not been presented for a degree in any other university, and all sources of materials used for this dissertation have been duly acknowledged.

Name: Miheretu Fufa Geleta

Signature: _____

Date: _____

DEDICATION

This thesis is dedicated to my lovely wife, Wubayehu Tesfa, in recognition of her support, love, and encouragement, which she so generously gives to me. I also dedicate it to my children Oliti, Letera, Ayantu, Bontu, and Tolera Miheretu, who endured the pain of separation and did not get much attention and love from me, during this study period.

BIOGRAPHICAL SKETCH

The author, Miheretu Fufa Geleta was born in Wellega, Ethiopia. He attended his Elementary School at Bururi Mariam Elementary School, his 6th and 7th grade at Kakie Primary School, and his Secondary School at Kellem Comprensive Secondary School. Then he joined Haramaya University, the then Alemaya University, in 1997 and graduated with BSc. degree in Plant Sciences in 2001. After his graduation, he was employed as a junior researcher by Oromia Agricultural Research Institute at Sinana Agricultural Research Center. After serving for seven years, he joined Jimma University for MSc study and graduated with MSc. degree in Plant Biotechnology in 2012. After his graduation, he re-joined Sinana Agricultural Research Center and served for about eight months before he got transfer to Adami Tulu Agricultural Research Center in 2013. He served at different research position and published more than 13 scientific papers. He also served the Research Center as Center Director for three years. Then, he joined Hawassa University in October, 2019 to pursue his PhD degree in Plant Biotechnology.

ACKNOWLEDGEMENTS

First of all I give many thanks to God, who made everything possible right from the start to the end of my study. Nothing was done without the will of God. It is my pleasure to express my sincere thanks and deep sense of gratitude to my research advisors, Dr. Andargachew Gedebo from Hawassa University College of Agriculture, Hawassa University and Dr. Tesfaye Letta from Oromia Agricultural Research Institute for their kind encouragement, guidance, and constructive suggestions to design and implement the research work. My special thanks goes to Dr. Tesfaye Letta for his understanding, kind-heartedness, and keen support right from the initiation of my research until its implementation by allocating the budget required for both field work and genotyping-by-sequencing, which would not have been possible. I also extend my sincere gratitude to Dr. Dangachew Lule from Oromia Agricultural Research Institute for his kind support, encouragement, constructive comments, and guidance during my research work.

It is a pleasure to give my special appreciation to my wife, Wubayehu Tesfa; my kids, Oliti, Letera, Ayantu, Bontu, and Tolera Miheretu for their understanding, patience, encouragement, and support during my study period. I appreciate Lencho Tesfa and Lalisa Tsegaye for their kind support during my study period in the preparation of seeds for planting, planting, and data collection. I also extend my appreciation to my Father, Fufa Geleta and my mother, Bushute Itana, for their support, encouragement, and initiation to pursue my education right from elementary until my PhD study. My father's devotion in educating me has laid the foundation for my today's success.

I am grateful to the staff of Adami Tulu Agricultural Research Center particularly Kasim Gudeta, Fatuma Sefawo, Bekelech Akele, Kufa Edao, Tsehay Ahmed, Megersa Gemmechu, Abdella Dadi, Mohammad Geleto, Gerema Amente, Dejene Kefena, Kasahun Kitila, Yasin Esmael, and Bekele Kedir for their keen support in my research work. I also thank Teshome Bogale, Director General of Oromia Agricultural Research Institute, Mr. Tafa Jobie, Crop Research Directorate Director, and Dr. Lemma Abera, Deputy Director General of Oromia Agricultural Research Institute, for their decision and support in the financing of my genotyping service, without which this study couldn't been possible.

Staff members of Sinana Agricultural Research Center and Debrezeit Agricultural Research Center in general; Tilahun Bayisa, Tesfaye Tadesse, Siraj Defo, and Demisie

Niguse in particular, deserve special thanks for their support on field experimentation during the study period.

ABBREVIATIONS AND ACRONYMS

AMOVA Analysis of Molecular Variance

ANOVA Analysis of Variance

BLINK Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway

BLUE Best Linear Unbiased Estimate

DArT Diversity Array Technology

DZARC Debrezeit Agricultural Research Centre

GA Genetic Advance

GAM Genetic Advance as a Percentage of Mean

GAPIT Genome Association and Prediction Integrated Tools

GBS Genotyping-by-Sequencing

GCV Genotypic Coefficient of Variation

GWAS Genome Wide Association Study

H^2 Broad Sense Heritability

IBC Institute of Biodiversity Conservation

IBPGR Institute of Biodiversity Plant Genetic Research

PCA Principal Component analysis

PCV Phenotypic Coefficient of Variation

PIC Polymorphic Information Content

LD Linkage Disequilibrium

MAF Minor Allele Frequency

Mbp Mega base pair

MTA Marker Traits Association

Q-Q Quantile -Quantile

QTL Quantitative Trait Locus

SARC Sinana Agricultural Research Centre

SNP Single Nucleotide Polymorphism

TASSEL Trait Analysis by Association, Evolution and Linkage

TITLES of PAPERS/MANUSCRIPTS

Chapter II

Genetic Diversity in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia based on phenotypic traits published as: Miheretu Fufa, Andargachew Gedebo, Tesfaye Letta & Dangachew Lule Genetic variation, genetic advance, heritability and correlation analysis of phenotypic traits in tetraploid wheat (*Triticum turgidum* spp.) landraces and some improved cultivars of Ethiopia Plant Genetic Resources 22(1):1-9 DOI: 10.1017/S1479262123001089.

Chapter III

Genetic Diversity of Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia based on grain quality traits

Chapter IV

Genetic Diversity in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia based on DArTSeq markers published as: Miheretu Fufa, Andargachew Gedebo, Tesfaye Letta & Dangachew Lule Genetic diversity and population structure analysis in tetraploid wheat (*Triticum turgidum* spp) germpalsm from Ethiopia based on DArTSeq markers *Genet Resour Crop Evol* (2023). https://doi.org/10.1007/s10722-023-01775-6.

Chapter V

Genome-Wide Association Mapping in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia

TABLES OF CONTENTS

Contents **Page**
ADVISORS' APPROVAL SHEET ii
EXAMINERS' APPROVAL SHEET iii
DECLARATION iv
DEDICATION v
BIOGRAPHICAL SKETCH vi
ACKNOWLEDGEMENTS vii
ABBREVIATIONS AND ACRONYMS ix
TITLES of PAPERS/MANUSCRIPTS xi
TABLES OF CONTENTS xii
LIST OF TABLES xv
LIST OF FIGURES xvii
LIST OF APPENDICE xix
Thesis Abstract xx
CHAPTER I 1
1. GENERAL INTRODUCTION 1
1.1. Background and Justification 1
1.2. Statement of the problems 3
1.3. Objectives of the study: 6
1.3.1. Major Objective 6
1.3.2. Specific Objectives 6
1.4. Research questions 6
1.5. Research hypothesis 6
1.6. Significance of the study 7
1.7. Thesis Organization 7
CHAPTER II 8
2. Literature Review 8
2.1. Origin, Evolution, and Genome of tetraploid wheat 8
2.2. Domestication of tetraploid wheat 10
2.3. Genetic Diversity of tetraploid wheat 12
2.5. Wheat Grain Quality traits 14
2.5.1. Protein content 15
2.5.2. Gluten strength 15
2.5.3. Vitreousness of wheat 16
2.6. Molecular markers and their applications in tetraploid wheat 17

2.7.1. Linkage based QTL mapping 19
2.7.2. Association mapping 20
2.8. Approaches for association mapping 21
2.8.2. Candidate gene association mapping 22
CHAPTER III 24
3. Genetic Diversity of Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm from Ethiopia based on Phenotypic Traits 24
Abstract 24
3.1. Introduction 25
3.2. Materials and Methods 28
3.2.1. Description of the study areas 28
3.2.2. Planting materials and experimental procedures 28
3.2.3. Qualitative morphological traits data collection and analysis 29
3.2.4. Quantitative morphological traits data collection 30
3.2.5. Quantitative data analysis 31
3.3. Results and Discussion 36
3.3.1. Qualitative traits diversity 36
3.3.2. Quantitative traits analysis 43
3.5. Conclusion 71
CHAPTER IV 72
4. Genetic Diversity in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia based on grain quality traits 72
4.1. Introduction 73
4.2. Material and Methods 75
4.2.1. Planting material 75
4.2.2. Grain quality traits data collection and analysis 75
4.2.3. Grain quality traits data analysis 75
4.3. Results and Discussion 79
4.3.1. Analysis of variance 79
4.3.2. Performance and patterns of variation of grain quality traits of the genotypes 80
4.3.4. Phenotypic and genotypic coefficient of variation 88
4.3.5. Estimate of broad sense heritability (H^2) and genetic advance 89
4.3.6. Correlation analysis 91
4.3.7. Principal Component analysis 92
4.3.8. Cluster analysis 97
4.5. Conclusion 103
CHAPTER V 104

5. Genetic diversity in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia based on DArTSeq markers 104
5.1. Introduction 105
5.2. Materials and Methods 107
5.2.1. Planting materials 107
5.2.2. DNA extraction and genotyping 107
5.2.3. Quality control, SNP calling, and filtering 107
5.2.4. Population structure and genetic diversity analysis 108
5.2.5. Analysis of molecular variance (AMOVA) and genetic diversity indices 109
5.3. Results and Discussion 110
5.3.1. SNP markers distributions 110
5.3.2. Population structure and genetic relationships 114
5.3.3. Principal component analysis 122
5.3.4. Genetic diversity indices analysis 124
5.3.5. Allelic patterns across the populations of tetraploid wheat 125
5.3.6. Analysis of molecular variance 134
5.3.6. Cluster analysis 138
5.5. Conclusion 141
CHAPTER VI 142
6. Genome-Wide Association Study in Tetraploid Wheat (*Triticum turgidum* spp.) Germplasm of Ethiopia 142
6.1. Introduction 143
6.2. Materials and Methods 145
6.3. Results and Discussion 147
6.3.1. Combined ANOVA 147
6.3.3. Linkage disequilibrium: 151
6.3.4. Marker Traits Association (MTA) 153
6.3.5. Pleiotropic locus 162
6.3.6. Functional annotation 163
6.4. Conclusion 165
CHAPTER VII 166
7. Summary and Recommendations 166
8. References 175

LIST OF TABLES

Table 3. 1 Characters used for the study, their codes, and descriptions 30

Table 3. 2 Mean squares and Estimates of the Shannon-Weaver diversity Index (H') for phenotypic traits in 174 landraces based on altitudinal class, regions of collections, and species 38

Table 3. 3 Estimates of the Shannon-Weaver diversity index (H') according to the within and between tetraploid wheat regions of collections and altitudinal classes for the qualitative traits assessed on 174 landraces 42

Table 3. 4 Mean squares of combined ANOVA of phenotypic traits of 196 tetraploid wheat accessions evaluated at Sinana and Debrezeit during 2020 45

Table 3. 5 The mean, minimum, maximum, and range values of quantitative traits at the entire genotypic level. 47

Table 4. 1 Mean squares of grain quality traits from combined ANOVA of 196 tetraploid wheat genotypes from Ethiopia 80

Table 4. 2 Range, mean, minimum, and maximum of grain quality traits of the 174 landraces along with 22 improved varieties 81

Table 4. 3 Selected tetraploid wheat genotypes based on grain quality traits performance 83

Table 4. 4 Comparison of mean performance of top 5% tetraploid wheat genotypes with mean of the population and released varieties for grain quality traits of tetraploid wheat grown at Sinana and Debrezeit during 2020 87

Table 4. 5 Comparison of mean performances of selected top 5% genotypes with the population mean performances for grain quality traits of tetraploid wheat germplasm grown at Sinana and Debrezeit during 2020 88

Table 4. 6 Variability, heritability, and genetic advance of tetraploid wheat germplasm based on grain quality traits 90

Table 4. 7 Pearson's correlation analysis for the eight grain quality traits of 174 landraces along with 22 improved varieties 91

Table 4. 8 Principal component analysis for the eight grain quality traits of 174 landraces along with 22 improved varieties 93

Table 4. 9 Principal component analysis for the eight grain quality traits of tetraploid wheat based on region of collection, altitudinal, and species 96

Table 4. 10 Clustering of 174 tetraploid wheat landraces along with 22 improved cultivars into two groups using eight grain quality traits 98

Table 4. 11 Intra- and inter-cluster distances of tetraploid wheat germplasm based on grain quality traits using average group linkage produced from centroid 99

Table 4. 12 Euclidean Distance of regions of collections of 174 tetraploid wheat landraces along with 22 improved cultivars based on grain quality traits 100

Table 4. 13 Euclidean Distance of altitudinal classes of 174 tetraploid wheat landraces along with 22 improved cultivars based on grain quality traits. 101

Table 4. 14 Euclidean Distance of species of 174 tetraploid wheat landraces along with 22 improved cultivars based on grain quality traits 102

Table 5. 1 The mean values of Minor allele frequency, Gene diversity, heterozygosity, and PIC across chromosome of tetraploid wheat study panel, at species level, regions of collection, and the whole population level 113

Table 5. 2 Inference of the optimal numbers of clusters existed in Ethiopian tetraploid wheat panel using Delta K Statistics based on DArTSeq markers 117

Table 5. 3 Population structure values generated using DArTSeq markers for study panel and landraces of tetraploid wheat 119

Table 5. 4 Inference of the optimal numbers of clusters existed in Ethiopian tetraploid wheat landraces panel using Delta K Statistics based on DArTSeq markers 121

Table 5. 5 Mean of different genetic indices of the genetic parameters in populations of tetraploid wheat genotypes 125

Table 5. 6 Total and mean allelic patterns across populations of the study panel and the landraces based on DArTSeq markers 128

Table 5. 7 Total and mean allelic patterns across the species, regions of collections, and altitudinal classes of tetraploid wheat of Ethiopia based on DArTSeq markers 132

Table 5. 8 AMOVA of tetraploid wheat germplasm at study panel, landraces, regions of collections, species, and altitudinal class level 137

Table 5. 9 Euclidean distances among the populations of tetraploid wheat genotypes at species, regions of collections, and altitudinal classes 140

Table 6. 1 Mean squares from combined ANOVA of phenotypic and grain quality traits showing significant MTA in tetraploid wheat of Ethiopia 148

Table 6. 2 Summary of linkage disequilibrium analyses among marker pairs and the number of significant marker pairs per chromosome and genome 151

Table 6. 3 GWAS result obtained using GAPIT for phenotypic and grain quality traits showing significant marker trait association 154

LIST OF FIGURES

Figure 2. 1 The evolution of wheat from the prehistoric Stone Age grasses to modern macaroni wheat and bread wheat 9

Figure 3. 1 Geographical map of Ethiopia, indicating areas of collection of the landraces 29

Figure 3. 2 Major phenotypic classes of qualitative traits recorded in tetraploid wheat landraces and improved cultivars of Ethiopia. 39

Figure 3. 3 Clinal variation in qualitative traits of Ethiopian tetraploid wheat germplasm 41

Figure 3. 4 The number of superior and inferior tetraploid wheat genotypes over (A) the population means, (B) the mean of released varieties, and (C) comparison of number of superior genotypes over population mean and released varieties' mean of phenotypic traits. 49

Figure 3. 5 PCA Biplot of whole genotypes of tetraploid wheat of Ethiopia based on 15 phenotypic traits 60

Figure 3. 6 PCA- Biplot of regions of collection of tetraploid wheat genotypes based on the fifteen phenotypic traits.. 62

Figure 3. 7 Dendrogram showing the clustering pattern of the 174 tetraploid wheat landraces along with 22 improved cultivars based on the fifteen phenotypic traits 67

Figure 3. 8 Dendrogram of the regions of collections of 174 landraces along with 22 improved varieties based on 15 phenotypic traits. 68

Figure 3. 9 Dendrogram of species of 174 landraces along with 22 improved varieties based on 15 phenotypic traits 69

Figure 3. 10 Dendrogram of altitudinal classes of 174 landraces based on 15 phenotypic traits 70

Figure 4. 1 The number of superior and inferior tetraploid wheat genotypes over (A) the population means and (B) the mean of released varieties grown during 2020 at Sinana and Debrezeit based on grain quality traits. 85

Figure 4. 2 Principal component analysis of eight grain quality traits of174 tetraploid wheat along with 22 improved varieties. 93

Figure 4. 3 PCA of grain quality traits of tetraploid wheat based on regions of collections (left), altitudinal (middle), and species (right) 94

Figure 4. 4 Dendrogram showing the clustering pattern of 174 tetraploid wheat landraces along with 22 improved cultivars based on the eight grain quality traits 99

Figure 4. 5 Dendrogram showing the clustering pattern of the regions of collection of Ethiopian tetraploid wheat germplasm based on grain quality traits 100

Figure 4. 6 Dendrogram showing the clustering pattern of the altitudinal classes of Ethiopian tetraploid wheat germplasm based on grain quality traits 101

Figure 4. 7 Dendrogram showing the clustering pattern of the species of Ethiopian tetraploid wheat germplasm based on grain quality traits 102

Figure 5. 1 Distribution of 10349 Single nucleotide polymorphisms (SNPs) across the Tetraploid wheat genome. Genome A and B are marked with blue and red colours, respectively 112

Figure 5. 2 Minor allele frequency distribution of tetraploid wheat study panel before and after filtering 112

Figure 5. 3 Population structure analysis of tetraploid wheat study panel based on DArTSeq markers 116

Figure 5. 4 Neighbor Joining (NJ) tree generated from Euclidean distance based on Ward.D method using 2381 DArTSeq markers from 184 tetraploid wheat genotypes. 116

Figure 5. 5 Principal component and genetic relatedness of 184 tetraploid wheat genotypes based on 2,381 genome-wide scanned, high-quality SNPs.. 118

Figure 5. 6 Population structure analysis of tetraploid wheat landraces based on DArTSeq markers. 120

Figure 5. 7 Neighbor Joining (NJ) tree generated from Euclidean distance based on Ward.D method using 2381 DArTSeq markers from 162 tetraploid wheat landraces. 121

Figure 5. 8 Population structure analysis of 184 tetraploid wheat genotypes based on the principal component analysis clustering as revealed by the first two principal components (PC1 and PC2) produced using 2381 polymorphic DArTseq markers 123

Figure 5. 9 Population structure analysis of 162 tetraploid wheat landraces based on the principal component analysis clustering as revealed by the first two principal components (PC1 and PC2) produced using 2381 polymorphic DArTseq markers 124

Figure 5. 10 Allelic patterns across the populations of tetraploid wheat study panel and landraces based on DArTSeq markers. 126

Figure 5. 11 Allelic patterns across the populations of tetraploid wheat germplasm at regions of collection, species level, and altitudinal classes based on DArTSeq markers. 130

Figure 5. 12 Cluster analysis of tetraploid wheat genotypes based on regions of collections, species level, and altitudinal classes using DArTSeq markers. 139

Figure 6. 1 Frequency distribution of traits having significant marker trait association in 184 tetraploid wheat germplasm of Ethiopia 150

Figure 6. 2 The plot of decay in linkage disequilibrium 152

Figure 6. 3 The distribution of significant marker pairs across the genomes of tetraploid wheat 152

Figure 6. 4 Manhattan plots and Q-Q plots for the GWA scan on phenotypic traits and grain quality traits.. 158

Figure 6. 5 Genomic positions of detected putative QTLs associated with phenotypic and grain quality traits. 161

LIST OF APPENDICE

Annex 1 seed source, pedigree, and passport data of tetraploid wheat of Ethiopia germplasm used in the genetic diversity study 196

Annex 2 The percentage frequency of four qualitative traits of 174 tetraploid wheat landraces along with the 22 improved varieties based on the regions of collections, the species level, and the altitudinal classes 200

Annex 3 Mean values of phenotypic and grain quality traits of 196 genotypes evaluated at Sinana and Debrezeit during 2020 202

Annex 4 Summary of the putative QTLs identified across tetraploid wheat chromosomes for agronomic and grain quality traits combined across locations and years 209

Annex 5 Functional annotation of genes identified within the QTL regions 211

Thesis Abstract

Genetic Diversity and Genome-Wide Association Mapping in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia

PhD Dissertation

Miheretu Fufa Geleta

Hawassa University, 2024

Ethiopia is a center of diversity for Tetraploid wheat (Triticum turgidum spp.) species, which hold ample genetic variation; howevr, the country remains a net importer of wheat due to a huge gap between production and consumption. The present study was aimed to investigate the extent and pattern of diversity based on phenotypic traits, grain quality traits, and molecular markers; and Genome-Wide association studies for phenotypic and grain quality traits.

Diversity assessment at the entire genotypes based on the qualitative traits revealed intermediate (0.60±0.01) diversity and high diversity concerning glum color (0.78), spike density (0.61), and seed color (0.86); however, low diversity for awndness (0.15). All qualitative traits showed highly significant ($p<0.001$) variation across regions of collections and altitudinal classes; however, only spike density was significantly ($p<0.05$) different at the species level. Within populations diversity was higher than between population diversity. Analysis of variance based on the quantitative traits revealed highly significant variation ($p<0.001$) among genotypes and for genotype by location interaction in all traits except for the number of effective tillers per plant. The observed mean and range values of the phenotypic traits revealed high variability in the accessions. Phenotypic Coefficient of Variation (PCV) and Genotypic Coefficient of Variation (GCV) values were high for grain yield, biomass yield, and harvest index. The estimates of heritability (H^2) ranged from 41.78 to 84.62 % respectively for grain yield and the number of seeds per spikelet. High genetic advance as a percentage of mean was observed for the number of seeds per spikelet, the number of effective tillers per plant, grain yield, biomass yield, and harvest index. The number of seeds per spikelet gave a high value of genetic advance and heritability implying high genetic gain from its selection. Grain yield showed a highly significant ($p<0.001$) negative correlation with days to booting and days to maturity and a positive correlation with the remaining quantitative traits.

*The combined ANOVA based on the grain quality traits revealed a highly significant variation ($p < 0.001$) among the genotypes and for genotype by location interaction for all traits. The genotypes showed a wide range in their thousand kernel weight, zeleny index, and vitreousness percentage. Grain yield and protein content, respectively, showed high and low values of PCV and GCV; however, an intermediate value of PCV and GCV was observed for gluten content and thousand kernel weight. Zeleny index and vitreousness percentage showed high PCV and intermediate GCV; however, protein content showed low PCV and GCV. The estimate of broad sense heritability (H^2) ranged from 33.57 % for vitreousness percentage to 66.36 % for zeleny index. The values of genetic advance and GAM observed were respectively high for grain yield (31.94, 26.91) and zeleny index (31.89, 29.34); intermediate for thousand kernel weight (18.27, 16.59), gluten content (15.70, 13.55), and vitreousness percentage (15.75, 12.97) and low for protein content (9.81, 8.44). Grain yield showed a highly significant positive correlation with thousand kernel weight (0.43***) and vitreousness percentage (0.19**) and positive correlation, but not significant, with protein content, gluten content, and zeleny index.*

Principal component analysis (PCA) based on phenotypic traits showed that the first and second principal components (PC) respectively accounted for 19.74% and 15.96 of the total variation in the entire genotypes, respectively. The first five principal components combined explained 61.21 % based on the entire genotypes. On the other hand, PCA of grain quality traits showed that the first and second PCs respectively accounted for 34.12 % and 20.65% of the total variation and the three PCs explained 69.11% of the total variation in the entire genotypes. Eleven landraces and eleven improved varieties were top performing based on their grain yield and grain quality traits. Clustering based on phenotypic traits and grain quality traits respectively grouped the whole genotypes into seven and six clusters without clear regional grouping.
Furthermore, the genetic diversity and population structure were investigated for the study panel as well as the landraces based on 10349 DArTSeq markers distributed across A and B genomes. The study panel was clustered into two populations: pop1 and pop2 respectively comprising 150 landraces and 3 released varieties and 19 released varieties and 12 landraces pop2 being more diverse than pop1; likewise, landraces were clustered into two subgroups. Analysis of molecular variance (AMOVA) revealed highly significant ($P<0.001$) variation between populations and within populations the latter being higher than the former implying that more attention should be given to individual accessions within populations to explore the existing genetic diversity. Little to moderate genetic differentiation (Fst) was respectively reported for the landraces (0.028) and the study panel (0.117) implying no significant differentiation among populations. High gene flow (1.297-8.818) was recorded based on DArTSeq markers at all level.

PCA based on DArTSeq markers grouped the released varieties and the landraces separately in the study panel; however, PCA of the landraces did not show clear groupings. 127 genotypes were reported to have one or more private alleles at 755 loci, an indication of key adaptive genes at these loci to be exploited in breeding programs. In conclusion, high genetic diversity was detected in Ethiopian tetraploid wheat germplasm, which could be utilized for future wheat breeding programs. The landraces ETW115 and ETW135 and the released varieties (Tesfaye and Bekelcha) could be used in a crossing program owing to their maximum genetic distance.

Analysis of genome-wide LD in the whole collection showed that 44.98% of the total marker pairs had a significant LD ($p<0.01$) with a mean r^2 value of 0.129791. 44.59% (18668) and 8.62% (3609) of the significant marker pairs showed r^2 values above 0.2 and 0.7, respectively, indicating a higher level of LD in the genome. The number of significant marker pairs observed was higher in the B genome (24.44%) than the A genome (20.54%). GWAS identified 44 SNPs, two of which are pleiotropic, across chromosome 1A, 2A, 2B, 3A, 3B, 4B, 5A, 5B, 6A, 6B, and 7B that were significantly (FDR <0.05) associated with some of the traits studied. Thirty-five QTLs were reported while the remaining 9 QTLs were found to be novel, which need to be validated for further use in the breeding program. We also identified several candidate genes, potentially regulating the traits, and encoding various proteins involved in plant growth and development. Hence, this study highlights the significance of the Ethiopian tetraploid wheat gene pool for improving tetraploid wheat globally; thus, a breeding strategy focusing on accumulating favorable alleles at these loci could improve tetraploid wheat production in Ethiopia and beyond.

Keywords: DArTSeq; Genetic diversity; Grain quality traits; GWAS; Heritability; Landraces; MTA; Phenotypic traits; Population structure; Tetraploid wheat

CHAPTER I

1. GENERAL INTRODUCTION

1.1. Background and Justification

There is a pressing need to increase production to feed the ever-increasing world population, which is estimated to reach almost 10 billion by 2050 (FAO 2017; FAO 2019; FAO 2011). Thus, agriculture, key to poverty and hunger alleviation in rural areas, is expected to produce much more to meet the need (FAO 2017). Likewise, in Ethiopia, agriculture is key driver of economic growth and food security for it directly offers livelihood for 84% of the population and accounts for more than 42% of national GDP and the lion's share of foreign currency earnings of the country (IBC 2012). The only option to increase agricultural production is to intensify production of crops (Gibbon 2012). Globally as well as in sub-Saharan Africa including Ethiopia, wheat is a commodity given priority to attain the objective of food self-sufficiency and food and nutrition security (FAO 2017; FAO 2011; FAO 2019; Teklu and Hammer 2008; Alemu 2020; Nigus *et al*., 2022).

Wheat is a commodity with a high market value that generates income for farmers in Ethiopia, the largest wheat producer in sub-Saharan Africa. Two types of wheat are predominantly grown in Ethiopia: tetraploid wheat, an indigenous to the country, and hexaploid wheat, a recent introduction to the country (Alamerew *et al*., 2004; Getachew and Worede 1991).

Ethiopia is a center of diversity for the cultivated tetraploid (2n = 4X= 28, AABB) wheat (*Triticum turgidum* species) which hosts an exceptional wealth of genes (Haile *et al*., 2013) and is where a remarkable diversity of varieties of wheat is concentrated, the most important bread grain of the earth (Vavilov 1950). Tetraploid wheat (*Triticum turgidum* spp.) is an important species within the genus Triticum and has amount of genetic variation for many desirable agronomic traits and other breeding attributes and can be used as a parent of hexaploid wheat to improve its yield and quality. It is the second cultivated wheat after *Triticum aestivum* (Li *et al*., 2020; Bechere *et al*., 1996). Tetraploid wheat has been grown in Ethiopia for thousands of years; the great majority of cultivars grown being

landraces with distinct morphotypes (Haile *et al*., 2013; Bechere *et al*., 1996). For millennia, farmers prefer to grow mixtures of tetraploid wheat landraces to add variety to their diet and to reduce the risks of losses due to new disease or pest outbreak or due to unusual environmental conditions (Pecetti and Damania 1996). During its long period of cultivation in Ethiopia, tetraploid wheat species has acquired diverse sets of characteristics making Ethiopia a center of diversity for the crop (Haile *et al*., 2013).

The consumption and demand of wheat showed increasing trend (Gebreselassie, Haile, and Kalkuhl 2017). Generally in Africa and particularly in sub-Saharan Africa including Ethiopia, the demand for wheat is increasing faster than for any other food crops (Alemu *et al*., 2019) and continues to increase by higher percent due to the rapid population growth, increased urbanization, change in food habits (Zegeye *et al*., 2020; Tadesse *et al*., 2018), and low production and productivity (Alemu *et al*., 2019). Particularly, in Ethiopia the emergence of many agro-processing industries additionally contributed to the increasing demand for wheat products such as bread, pasta, macaroni, and confectionaries from time to time (Zegeye *et al*., 2020). The low yield of and increased demand for wheat will be a major challenge, particularly in developing countries, where urban population growth is forecasted to increase by higher percent (Alemu *et al*., 2019). Wheat is the grain imported with large volume in Ethiopia (Gebreselassie *et al*., 2017) to fill the gap between production and consumption. Thus, enhancing wheat yield is one of the long-term solutions to overcome wheat food problems and to meet the demand of the growing population of the country (Zegeye *et al*., 2020). Following several government programs and initiatives implemented to drive agricultural growth and food security in the country (Gebreselassie *et al*., 2017) and research effort so far made in Ethiopia, wheat production enhanced significantly with the annual wheat production of Ethiopia being about 5.8 million tons with mean productivity of 3 tons per hectare (tha−1)(CSA 2021). Despite the progress so far made in enhancing wheat production, the current mean productivity of wheat is below global level (Alemu *et al*., 2019) and below the attainable yield of the crop which reaches up to 5t/ha (Nigus *et al*., 2022; Zegeye *et al*., 2020). Hence, low productivity continued to be the major challenge facing wheat (Alemu *et al*., 2019) as it could not cope with the ever-increasing demand. Accordingly, enhancing wheat yield is a key part of the agenda in the Ethiopian government's food security policy programs (Gebreselassie *et al*., 2017) and also the focus of research in Ethiopia (EAIR 2017).

Moreover, to feed the ever-increasing global population, the current production needs to increase to 10 million tons in 2050 (Zegeye *et al*., 2020).

1.2. Statement of the problems

Ethiopian tetraploid wheat landraces are endowed with valuable wealth of genetic diversity which are (Zegeye *et al*., 2020; Haile *et al*., 2013; Muleta *et al*., 2017) highly prized for their potential value as source of variation in crop improvement programs for different economically important traits (Liu *et al*., 2017; Mengistu and Pè 2016; Hardon 1992) such as earliness, disease and pest resistance, nutritional quality, resistance to drought and other stress conditions, and characteristics useful in low-input agriculture (Hardon 1992). Thus, they can be used as parents to improve yield and quality of *Triticum aestivum* in breeding program (Li *et al*., 2020). Such a rich source of genetic diversity provides a broad genetic foundation for plant breeding and genetic research (Thapa *et al*., 2015). More than 7000 tetraploid wheat landraces collections were found in Ethiopian Biodiversity Institute (Negisho *et al*., 2021). However, only limited portions of them were exploited in the modern breeding efforts (Liu *et al*., 2017; Mengistu and Pè 2016). For example, only 2% of the improved durum wheat varieties cultivated in Ethiopia were composed gene from Ethiopian landraces. The remaining 98% is introduction from exotic materials. This is associated with lack of comprehensive characterization of the landraces for important quantitative and qualitative traits such as disease resistance (Mengistu and Pè 2016). Therefore, assessing the genetic diversity of tetraploid wheat landraces is crucial to address the different breeding purposes such as stable and high grain yield, and grain quality. Knowledge of genetic diversity is very important to identify, conserve, and utilize useful materials from the existing landraces (Asmamaw *et al*., 2020), has a significant impact on the improvement of crop plants, and is vital to meet the diversified goal of breeding (Asmamaw *et al*., 2019). We can get a better knowledge of the collected material using morphological, biochemical, and molecular markers (Teklu and Hammer 2008; Govindaraj *et al*., 2015). Morphological markers have been in use for diversity analysis starting in the era of selection-based plant breeding and continue to play a role even today when molecular markers are being used. They are the alternative to their molecular counterparts for assessing genetic diversity. Highly heritable morphological characters are still applicable for germplasm characterization and management and can be analysed anywhere in the world without needing laboratory capacity, or the expense of sophisticated equipment and chemicals (Hailu 2011). Many of these morphological characters serve as

valuable markers in genetic analysis while others such as ear shape or awning may be of interest for their direct and indirect effects on the plant's yield potential (Blckman and Payne 1994). To unlock genetic diversity of tetraploid wheat of Ethiopia, attempts were made by many researchers using morphological diversity (Getachew and Worede 1991; Bechere *et al*., 1996; Eticha *et al*., 2005; Hailu *et al*., 2006; Belay *et al*., 1993,1997; Teklu and Hammer 2008; Alamerew *et al*., 2004; Mengistu *et al*., 2015).

In addition to enhancing yield of wheat, improving wheat grain quality is also of vital value owing to the current increase in the requirement for high quality food, especially for wheat-based food products (Bhalla and Singh 2017). Wheat grain quality is central to food processing and nutritional value and is a decisive factor for consumer acceptance and commercial value of wheat cultivars (Wang *et al*., 2017). Grain quality diversity was assessed by many authors using gluten strength and yellow pigment (Hailu and Merker 2008; Belay and Merker 1994), variation in high- and low-molecular weight glutenin subunits, and omega gliadins (Hailu *et al*., 2006) and some grain quality (protein, vitreousness, ash and moisture content) (Tsegaye *et al*., 2012). But, these are insufficient. Therefore, assessing the genetic diversity of the landraces based on their grain quality traits is of great value in developing cultivars with the required end use quality.

Furthermore, assessing genetic diversity of tetraploid wheat using molecular markers of paramount importance because they are not influenced by environmental conditions and reflect genetic variation more faithfully (Salvador-Figueroa *et al*., 2015). Numerous DNA-based genetic markers were developed and contributed greatly to our current understanding of genome organization and genetic variation, but they are constrained by their dependence on gel electrophoresis. To overcome such restrictions, hybridization-based methods using nucleic acids immobilised on solid-state surfaces have been developed. DNA chips were developed to analyse genotypes for single nucleotide polymorphisms (SNP) but the technology depends on intensive genomic sequencing and a high cost of analysis that cannot be matched in agriculture or basic research. A technology with similar power but with a much greater breadth and much lower cost for DNA polymorphism analysis called Diversity Array Technology (DArT) was reported. Genetic marker analysis through Diversity Arrays offers a low-cost high-throughput, robust system with minimal DNA sample requirement capable of providing comprehensive genome coverage even in organism without any DNA sequence information (Jaccoud *et al*., 2001). Diversity Arrays Technology (DArT) can detect any type of DNA variation at several hundred genomic loci

in parallel without relying on sequence information (Wenzl *et al*., 2016) and has been applied successfully for polyploidy and complex genomes such as tetraploid and hexaploid wheat (Robbana *et al*., 2019). DArTseq markers were widely applied for genetic diversity studies in many plant species (Ficco *et al*. 2012) such as wheat (Akbari *et al*. 2006); however, there is no or limited work in the GD study of Tetraploid wheat of Ethiopia.

In addition to assessing the genetic diversity, undertaking association mapping is of great value to detect associations between traits of interest and genetic markers (Hu *et al*., 2015). Association mapping, which exploits historical recombination events existing in the breeding populations being studied, is a powerful technique used to identify marker-trait associations (Bajgain *et al*., 2015). The key to modern plant breeding is elucidating the genetic basis of complex quantitative traits controlling agronomic traits of economic importance (Alemu *et al*., 2020) using recent genomic based research. This will facilitate the utilization of Ethiopian tetraploid germplasm in the crop improvement program. GWAS results in tetraploid wheat of Ethiopia were reported on yield and related traits, root system arctecture, disease resistance, grain shape and color revealing important insight on the association of traits of economic importance and genetic markers; however, it is limited and insufficient. DArTSeq has been widely and successfully applied for association mapping in polyploidy and complex genomes such as tetraploid wheat (Robbana *et al*., 2019); however, its application in AM of tetraploid wheat of Ethiopia is not available. Its application in the association mapping in tetraploid wheat of Ethiopia would reveal valuable alleles of economic importance.

The existing landraces collected and maintained by Ethiopian Biodiversity Institute (EBI) were not fully studied to effectively utilize them for different economically important traits. Only small portions of the huge collections at EBI were characterized (Negisho *et al*., 2021; Asmamaw *et al*., 2020; Alemu *et al*., 2020). Therefore, it is crucial to assess the genetic diversity of these landraces using morphological, industrial quality traits, and molecular markers. Additionally, detecting the association between traits of interest and genetic markers is of great value. Therefore, this project was initiated with the following objectives:

1.3. Objectives of the study:

1.3.1. Major Objective

To assess the genetic variation among tetraploid wheat landraces collected from different regions of Ethiopia along with released varieties using morphological, quality traits, and molecular markers and to identify elite genotypes with useful traits for future breeding purposes.

1.3.2. Specific Objectives

i. To assess the extent and pattern of diversity of tetraploid wheat germplasm of Ethiopia based on phenotypic traits

ii. To assess the extent and pattern of diversity of tetraploid wheat germplasm of Ethiopia based on grain quality traits

iii. To assess the genetic diversity and population structure of tetraploid wheat germplasm of Ethiopia based on molecular markers

iv. To study the association of phenotypic traits and grain quality traits with molecular markers

1.4. Research questions

So, this study aimed to answer the following research questions.

1. Are tetraploid wheat landraces diverse with respect to their morphology, grain quality traits and at molecular level? Or are they more diverse than the current improved cultivars under cultivation? What does the population structure of tetraploid wheat looks like?
2. Is there any relation or association between important traits and molecular markers?

1.5. Research hypothesis

Hypothesis 1: there is no distinct population structure and no diversity in the tetraploid wheat landraces and improved cultivars with respect to their morphological traits, grain quality traits.

Hypothesis 2: there is no distinct population structure and no diversity in the tetraploid wheat landraces and improved cultivars with respect to their grain quality traits.

Hypothesis 3: there is no distinct population structure and no diversity in the tetraploid wheat landraces and improved cultivars based on molecular markers.

Hypothesis 4: there is no relation between DArT markers and the important traits such as grain quality traits and some morphological traits.

1.6. Significance of the study

This research study contributed to the identification of some valuable germplasms with private allelels with key adaptive genes which are responsible for valuable grain quality traits and yield and yield related traits that can be used in future breeding programs.

1.7. Thesis Organization

A total of four activities were covered in this study to address the specific objectives stated. These include genetic diversity analysis using phenotypic traits, genetic diversity analysis using grain quality traits, genetic diversity analysis using DArTSeq markers, and genome wide association mapping. The study chapters are organized as follows:

Chapters	Title
Chapter 1	General Introduction
Chapter 2	Literature Review
Chapter 3	Genetic Diversity Based on Phenotypic Traits
Chapter 4	Genetic Diversity Based on Grain Quality Traits
Chapter 5	Genetic Diversity Based on DArTSeq Markers
Chapter 6	Genome Wide Association Study
Chapter 7	Summary and Recommendations

CHAPTER II

2. Literature Review

2.1. Origin, Evolution, and Genome of tetraploid wheat

Wheat is an ancient cultivated crop which is supposed to have originated in the Middle East (Morris C Peter 2000) and belongs to the Poaceae family, the largest family within the monocotyledonous plants (Chittaranjan Kole 2006). The family Poaceae (grasses) evolved 50– 70 million years ago (Mya) (Kellogg 2001; Huang *et al*., 2002) and the sub-family Pooideae including wheat, barley and oats diverged around 20 Mya (Inda *et al.*, 2008).

The key to understanding the evolution of wheat involves an elucidation of the evolution of the tetraploid wheat species (Carver 2009). The evolution of wild into domesticated emmer had occurred in cultivated fields over a period of several hundred years, following its initial cultivation, namely during the early PPNB (9,500–9,000 BP)(Feldman and Levy 2015) (Feldman and Kislev 2007).Tetraploid wheat (2n=28, BBAA) is one of the common present-day wheat cultivars derived from wild tetraploid progenitor T. dicoccoides (Özkan *et al*., 2011) which is a result of hybridization between two diploid progenitors classified in the genera Triticum (wild wheat) and Aegilops (the closest relative goat grass) as genome analysis studies revealed (Feldman and Levy 2015; Huang *et al*., 2002; Dvorak and Akhunov 2005). Strong evidence points to the wild outcrossing Ae. speltoides (SS) (or a genotype similar to it) as the female parent of tetraploid wheats and to wild T. urartu (A^uA^u) as the male parent (Kilian, Martin, and Salamini 2010). Wild diploid wheat (T. urartu, 2n = 2x = 14, genome A^uA^u) hybridized with the B genome ancestor that is the closest relative of goat grass (Aegilops speltoides, 2n = 2x = 14, genome SS) 300,000–500,000 years before present (BP) (Huang *et al*., 2002; Dvorak and Akhunov 2005) to produce wild emmer wheat (T. dicoccoides, 2n = 4x = 28, genome A^uA^uBB) (Figure 2.1).

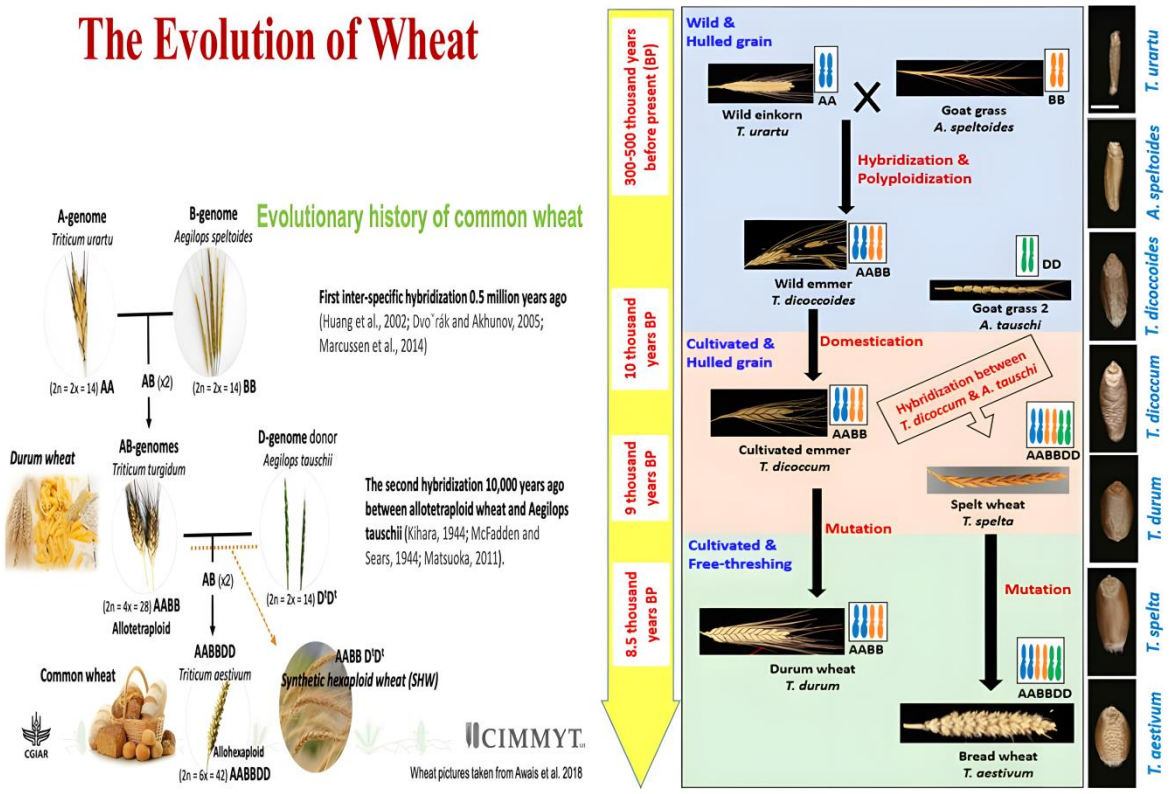


Figure 2. 1 The evolution of wheat from the prehistoric Stone Age grasses to modern macaroni wheat and bread wheat (adopted from González-Diéguez *et al*. 2022(left) and reproduced from https://www.mdpi.com/ijms/ijms-21-05836/article_deploy/html/images/ijms-21-05836-g001.png(right))

About 10,000 BP, hunter-gatherers began to cultivate wild emmer. Subconscious selection gradually created a cultivated emmer (T. dicoccum, 2n = 4x = 28, genome A^uA^uBB) that spontaneously hybridized with another goat grass (Ae. tauschii, 2n = 2x = 14, genome DD) around 9,000 BP to produce an early spelt (T. spelta, 2n = 6x = 42, genome A^uA^uBBDD) (http://www.newhallmill.org. uk/wht-evol.htm; Peng, Sun, and Nevo 2011). Over time, Emmer Wheat also mutated and evolved to create Durum Wheat, now used for making a wide range of pasta products such as macaroni, spaghetti and lasagna (http://www.newhallmill.org.uk/wht-evol.htm). The wild and domesticated forms of tetraploid wheats are genetically very close as observed from the high fertility of hybrid from a cross between wild emmer and *T. turgidum* ssp. *dicoccon* (Schrank) Thell. and ssp. *durum* (Feldman and Kislev 2007).

Tetraploid wheat contains one genome, A, that is derived from a diploid Triticum species and a second genome, B, from another, yet unknown diploid species. Hexaploid wheat contains the two genomes of tetraploid wheat, A and B, and a third genome, D, from

another species. Genome B, which is a modified genome, derived from Ae. speltoides or from a closely related species to Ae. speltoides, which is extinct or extant and underwent changes at the polyploid level (Feldman and Levy 2015). B genome has undergone significant intergenomic non-coded and coded DNA changes (in both the diploid and the tetraploid wheats) since the formation of tetraploid wheat. The B genome component of polyploid wheat is the largest of the wheat genomes and, because of the large degree of change at the DNA level, the true donor of the B genome since polyploid formation has been very difficult to establish. Regardless, since the cytoplasm donor is the female in the original cross creating the polyploid and is always listed first in any pedigree, the tetraploid genome designations should technically be BBAA or GGAA (Carver 2009). Genome A, that has been modified relatively little, was derived from T. urartu. Tests using DNA molecular markers show that wild tetraploid wheat did participate in the evolution of hexaploid wheat (J. Dvorak, Luo, and Akhunov 2011). Allohexaploid wheat (genome BBAADD) originated from hybridization between allotetraploid wheat and Ae. tauschii, the donor of the D genome. Modern classification for the genus Triticum recognizes two diploid species, T. monococcum L. and T. urartu Tum. ex Gand., two tetraploid species, T. turgidum L. and T. timopheevii (Zhuk.) Zhuk., and two hexaploid species, T. aestivum L. and T. zhukovskyi Men. & Er. The economically important wheat are T. aestivum (bread wheat, comprising 95 % of the global wheat production) and T. turgidum (macaroni wheat) (Feldman and Levy 2015).

2.2. Domestication of tetraploid wheat

Tetraploid emmer (*T. turgidum* ssp. dicoccum, genome AABB) and diploid einkorn (*T. monococcum*[i], genome AA) were the earliest domesticated forms of wheat (Heun *et al*., 1997). Wild emmer(*Triticum dicoccoides*) (2n = 4x = 28; genome BBAA) was first discovered in nature in 1906 in eastern Galilee, Israel, by Aaron Aaronsohn (Peng, Sun, and Nevo 2011) which significantly advanced the understanding the where, when, and how of wheat domestication (Feldman and Kislev 2007). It grows naturally all over the Fertile Crescent (Peng, Sun, and Nevo 2011) in a region extending from the southern Levant across Israel and Lebanon to south-eastern Turkey and across northern Iraq and north-western Iran. Archaeological record indicates that domesticated emmer first appeared in the southern Levant and in south-eastern Turkey about 9,500–9,000 BP (Jack R. Harlan and Zohary 1966).

The first cultivation of wheat occurred about 8,000 to 10,000 years ago in the Fertile Crescent (modern day Iraq and parts of Turkey, Syria and Iran), as part of the Neolithic Revolution when hunter-gatherer lifestyle transitioned into settled agriculture (Gopher, Abbo, and Lev-Yadun 2000; Salamini *et al*. 2002; Hu *et al*., 2016; Peng, Sun, and Nevo 2011). On the other hand, the first indication of cultivation of wild emmer in the southern Levantine Corridor was found in the Pre-Pottery Neolithic A (PPNA) (10,300–9,500 BP) (Feldman and Kislev 2007).Two wild tetraploid wheat species are known namely *T. dicoccoides* and *T. araraticum*. They are similar in morphology, but different in their genomic constitution: *T. dicoccoides*, has the genomic formula $BBA^{u}A^{u}$ and *T. araraticum* $GGA^{u}A^{u}$. *T. dicoccoides*, or wild emmer, belongs to the first cereals domesticated by humans in the Fertile Crescent and in its domesticated form is known as T. dicoccon (emmer, $BBA^{u}A^{u}$) (Kilian, Martin, and Salamini 2010). The domesticated form of T. dicoccoides is known as T. dicoccum (emmer, $A^{u}A^{u}BB$) (Peng, Sun, and Nevo 2011). Domesticated forms of emmer wheat (characterized by non-brittle spikes) appeared presumably at the beginning of the Pre-Pottery Neolithic B (PPNB; 9,500– 7,500 BP) (Feldman and Kislev 2007).

The process of domestication was mainly through selection of traits that separated them from their wild relatives (Shewry 2009). Two of the most important traits were the loss of shattering of the spike at maturity (determined by mutations at the Br (brittle rachis) locus), and the change from hulled forms (glumes adhere tightly to the grain), to free threshing naked forms (Nalam *et al*., 2006; Doebley, S.B, and Smith 2006). The development of free-threshing was due to mutation in the primitive q allele on chromosome 5A that led to the formation of the partially dominant Q allele, which modified the effects of recessive mutations at the Tg (tenacious glume) locus ((Jantasuriyarat *et al*., 2004; Simons *et al*., 2006; Dubcovsky and Dvorak 2007). The acquisition of a non-brittle rachis allowed early farmers to harvest the grain much more efficiently without premature loss of spikelets, while the development of free threshing forms allowed easy release of seeds from hulls during threshing. Other traits modified by the domestication process included larger seeds, loss of seed dormancy, and improved grain quality (Harlan, De Wet, and Price 1973).

2.3. Genetic Diversity of tetraploid wheat

Genetic diversity is the fundamental source of biodiversity. The improvement of cultivated plants considerably depends on the extent of genetic variability available within the species. The genetic variation that exists among plant populations is a basic requirement for efficient development and improvement of such populations. It also indicates whether a population can withstand and live with changes in the environment, which are mostly altered in an unpredictable way (Caliska 2012). Genetic diversity evaluations provide a great scientific basis for studies on species evolution, variety identification, parent selection, variety protection, and ultimately better utilization of germplasm and more efficient breeding (Lipka *et al*., 2012). Genetic diversity is important to a species' fitness, long-term viability, and ability to adapt to changing environmental conditions (Ahuja and Jain 2015)(Ahuja and Jain 2015) and provides opportunity to develop new and improved cultivars with desirable characteristics such as high yield, large seed, and pest and disease resistance (Govindaraj, Vetriventhan, and Srinivasan 2015). Ethiopia is considered to be one of the world's most important centres of crop diversity (Vavilov, 1950) including tetraploid wheat (Eticha *et al*., 2005). Vavilov (1929) was particularly impressed with the diversity of Ethiopian wheats. He made wide collection and recognized five species which include: 1. T. durum subsp. Abyssinicum, 2. T. turgidum subsp. Abyssinicum, 3. T. dicoccum, 4. T. vulgare (now T. aestivum hexaploid) and T. polonicum (Engels, Hawkes, and Worede 1991). Tetraploid wheat (2n = 4x = 28) have been under cultivation in Ethiopia for thousands of years, and have acquired a diverse set of characteristics that makes the country a centre of diversity for these species (Eticha *et al*., 2005). Tetraploid wheats are genetically and morphologically diverse and their evolution under domestication has not been fully elucidated (Oliveira *et al*., 2012). In particular, Ethiopian tetraploid wheat landraces hold extremely high potential of genetic diversity and are highly prized for their potential value as source of variation in crop improvement programs for different economically important traits (Liu *et al*., 2017; Mengistu and Pè 2016; Hardon 1992) such as earliness, disease and pest resistance, nutritional quality, resistance to drought and other stress conditions, characteristics useful in low-input agriculture (Hardon 1992) and stem rust resistance (Haile, Hammer, Ayele, *et al*., 2013). A rich source of genetic diversity provides a broad genetic foundation for plant breeding and genetic research (Thapa *et al*., 2015). More than 7000 tetraploid wheat landraces collections were found in Ethiopian Biodiversity Institute (Negisho *et al*., 2021). However, only limited

portions of them were exploited in the modern breeding efforts (Liu *et al*., 2017; Mengistu and Pè 2016). Only 2% of the improved durum wheat varieties cultivated in Ethiopia were composed gene from Ethiopian landraces. The remaining 98% is introduction from exotic materials (Mengistu and Pè 2016). This is associated with lack of comprehensive characterization of the landraces for important quantitative and qualitative traits such as disease resistance (Mengistu and Pè 2016). Therefore, assessing the genetic diversity of tetraploid wheat landraces is crucial to address the different breeding purposes such as stable and high grain yield, grain quality and rust resistance. Knowledge of genetic diversity is very important to identify, conserve and utilize useful materials from the existing landraces (Asmamaw, Keneni, and Tesfaye 2019) and has a significant impact on the improvement of crop plants. It helps us for efficient utilization of the genetic potential held in the landraces. Genetic diversity is the basis for genetic improvement (Huang *et al*., 2002). We can get a better knowledge of the collected material using morphological, biochemical and molecular markers (Teklu and Hammer 2008; Govindaraj, Vetriventhan, and Srinivasan 2015) to assess genetic diversity within and between populations. Morphological, biochemical and molecular markers can be employed in the assessment of genetic variability among genotypes within land races, populations and species (Eticha *et al*., 2005; Kidane *et al*., 2017).

2.4. Morphological markers of wheat

Morphological markers represent the expressions of plant phenotypic traits that are generated at a specific developmental stage or under a specific environmental condition. These markers have been widely used in the studies of many crops including wheat. However, morphological markers are largely limited by numbers and polymorphisms as well as some of their expressions are affected by gene expression regulation, developmental stage, and environmental factors. Thus, phenotypic differences often do not represent genotypic variations and time-consuming (Tian *et al*., 2015). Many of these morphological characters serve as valuable markers in genetic analysis while others such as ear shape or awning may be of interest for their direct and indirect effects on the plant's yield potential (Lupton, 1987). Hailu, Merker, Belay, *et al*. (2006) did multivariate analysis of diversity of tetraploid wheat from Ethiopia using twenty three morphological traits and found that plant height, days to heading, the length and density of the spike, and kernel colour were the most important traits contributing to variation. Belay *et al*. (1997)

also studied pattern of morphological diversity in thirty four tetraploid wheat landraces of Ethiopia using glume colour, glume pubescence, beak awn, seed colour and spike density and found that spike density was the only character showing difference between region and between altitudes. Bechere *et al*. (1996) also did phenotypic characterization on twenty seven populations of landraces from northern and north central regions of Ethiopia using glume colour, glume pubescence, spike density, spike length, seed colour, seed vitreousness, seed size, seed shape, beak length, spikelet per spike and seed per spike and found that only spike density showed significant difference.

2.5. Wheat Grain Quality traits

Grain quality is the second crop yield (Keler *et al.*, 2020). Currently, there is an increasing requirement for high quality food, especially for wheat-based food products (Bhalla and Singh 2017).With the increase in grain production, special attention is paid to improving its technological qualities such as the strength of flour, gluten quantity and quality, protein content, and others (Keler *et al.*, 2020). The nutritional quality of cereals is an important component of the world's diet as cereals represent the largest portion of world food supplies. More than other cereals, bread and durum wheat assume greater importance as a source of protein for much of the world's increasing population (Royo *et al*., 2005).

In addition to enhancing yield of wheat, improving wheat grain quality is also vital value. In the past, grain quality was not a strong criterion of variety selection. Only grain yield and disease resistance were given due attention in breeding program. However, things have changed as a result of changing food habits, increasing urbanization and trends towards raising middle class society. As a result, critically looking for better quality varieties suiting for preparation of different end products is becoming a criterion for the development of new varieties (Drikvand *et al*., 2018; Tadesse, Bishaw, and Assefa 2018). Additionally, Wheat grain quality is central to food processing and nutritional value, and is a decisive factor for consumer acceptance and commercial value of wheat cultivars (Wang *et al*., 2017). Quality characteristics including grain protein content and gluten content are among important contributors to the grain yield and quality of wheat (Drikvand *et al*., 2018; Tadesse, Bishaw, and Assefa 2018). Protein is generally accepted as a major indicator of physico-chemical properties of wheat flour (Sardana 2000).

Quality of wheat is judged by the grain protein content, gluten content, vireousness etc (Sardana, 2000). Three basic quality factors that determine wheat type are hardness, gluten

strength and protein content. In general, soft wheat with weak gluten and low protein content are not suitable for bread making; hard wheat with strong gluten and high-protein content are preferred (Bushuk and Rasper 1994). Wheat is classified in the market on the basis of kernel vitreousness and protein content (which in turn is an estimate of flour strength).

2.5.1. Protein content

Protein content is important in bread making because, other things being equal, high-protein flours have a high loaf volume potential, high water absorption and produce loaves with good keeping quality (Bushuk and Rasper, 1994). The relationship between crude protein content and bread making potential is so well established that the protein is generally accepted as a major indicator of physico-chemical properties of wheat flour (Sardana, 2000). Due largely to genetic selection, wheat of high protein content tend to be hard, to have strong gluten and to produce good bread, whereas wheat of low-protein content tend to be soft, to have weak gluten and to produce small loaves of inferior crumb structure (Bushuk and Rasper 1994).

A fairly wide range of protein content may be incorporated in bread flour grists, but wheat of less than 11 % protein are usually unsuitable for bread making if used alone. These type-determining factors are generally controlled genetically by variety although protein content may vary widely for a given variety according to location, soil fertility, rainfall, etc (Bushuk and Rasper 1994).

Grain protein content is a complex trait controlled by polygenic quantitative factors, with a low heritability, high environmental influence and usually negatively correlated with grain yield in most production environments. However, correlations are seldom sufficiently high to permit an increase of grain protein percentage by genetic manipulation without adversely depressing yield. More than other cereals, bread and durum wheat assume greater importance as a source of protein for much of the world's increasing population (Royo *et al*., 2005).

2.5.2. Gluten strength

Wheat is unique among cereal grains in that flour milled from it is capable of making a light, palatable, well-risen loaf of bread when processed into fermented dough. Gluten is the insoluble storage proteins of wheat that give wheat its functional properties. Gluten is

described as the most complicated rheological material known to man. It is chemically complex because it is made up of hundreds of different protein subunits which aggregate to form extremely large molecules when made into dough. It is gluten that gives dough its ability to form thin sheets that will stretch and hold gas. Even relatively poor quality wheat can produce bread that is significantly more palatable than that made from flour from other cereal grains (Bushuk and Rasper 1994). Gluten forms a cohesive mass, the main use of gluten in bread-making (Blackman and Payne 1994). In general, soft wheat with weak gluten and low protein content are not suitable for bread making; hard wheat with strong gluten and high-protein content are preferred (Bushuk and Rasper 1994). The higher the gluten content in flour and the better its quality is, the stronger the flour and the better the baking properties of flour are (Keler *et al*., 2020).

Doughs made from strong flours tend to be less viscous and more elastic than doughs from weak flours, but generally require more mixing time and energy in order to reach an optimum rheological state. Properly developed doughs have the ability to stretch and form thin sheets that will hold gas and promote production of light, well-risen loaves. Dough-handling properties during processing may range from soft and sticky for doughs from weak flours to extremely inelastic and tough for doughs produced from flours that are too strong. Well-balanced dough properties that fall between these two extremes are required for baking. Mixing requirements must also not be too short or too long (Bushuk and Rasper 1994).

2.5.3. Vitreousness of wheat

Vitreousness of wheat is a measure of its quality and is related to the protein content. Nonvitreous (starchy) kernels are opaque and softer, and result in decreased yield of coarse semolina. On the other hand, vitreous kernels appear hard, glassy and translucent, and have superior cooking quality and pasta color, along with coarser granulation and higher protein content. Thus, the vitreousness of wheat kernels is an important selection criterion in grain grading (Wang, Dowell, and Zhang 2002). Vitreousness, the starchy endosperm appearance, is an optical property attributed to differences in endosperm porosity. Grains are classified as mealy when the starchy endosperm is porous and appears white and floury, or as vitreous when it is transluscent and glassy (Chichti *et al*., 2018). Vitreosity, an optical property, is directly related to density through the relative porosity of the

endosperm. The presence of pores in the endosperm act to scatter light reflected from the endosperm, causing it to appear more opaque or mealy (Dobraszczyk *et al*., 2002).

2.6. Molecular markers and their applications in tetraploid wheat

Crop improvement through conventional breeding and genetics made significant contributions in the improvement of traits for disease resistance and environmental stresses. However, the progress in genetic gains is slow particularly in targeting traits with complex inheritance and addressing challenge brought by the changing climate, pathogen population and agricultural practices (Araus *et al*., 2008; Cooper *et al*., 2009). New breeding methods based on molecular marker technology offer great promise for plant breeding by providing solutions to overcome some of the constraints faced by classical phenotypic screening approaches in plant breeding. Molecular markers can enhance pyramiding of target traits into adapted cultivars with greater precision, reduced loss of genetic gain and shorter breeding cycles (Xu and Crouch 2008). Molecular markers, in a broad sense, refer to DNA sequences or biochemical markers (e.g., isozymes, allozymes, and some other protein-based markers) that are inheritable and detectable, while in a narrow sense, they are specific DNA sequences that can reflect the genetic differences among individuals or populations. Molecular markers are genetic markers that are derived from DNA nucleotide changes occurring in individuals of a given population and represent the genetic polymorphisms at the DNA level (Tian *et al*., 2015).

Molecular markers can be used in many areas of crop breeding programs (Wang, Dowell, and Zhang 2002) as a research tool to analyse existing populations (Bhalla and Singh 2017), for assessment of genetic diversity, germplasm identification, genetic linkage, map construction, map-based cloning, marker assisted breeding, gene mapping and for the mapping of complex traits (Nakaya and Isobe 2012; Meuwissen, Hayes, and Goddard 2001). Molecular markers are widely present in the genome. By comparing the polymorphisms of these randomly distributed genomic markers; one can evaluate the examined populations comprehensively, thereby revealing their genetic nature. These data can then be used in species/variety cluster analysis which will allow better understanding of their phylogenetic and genetic relationships. The high level of polymorphism of molecular markers will continue to play a very important role in kinship studies, species classification, germplasm identification, and protection (Tian *et al.*, 2015).

Various types of DNA-based molecular marker technologies were developed since the emergence of restriction fragment length polymorphisms (RFLPs) in the 1980s (Philips and Vasil 2001) and contributed greatly to our current understanding of genome organization and genetic variation (Jaccoud *et al*., 2001). However, they are limited in use mainly due to one or more of the following limitations: time-consuming multi-step protocols, low levels of polymorphisms, inability to detect single-base mutations, dominant, the use of radioactive isotopes (Somers, Isaac, and Edwards 2004) and dependence on gel electrophoresis (Jaccoud *et al*., 2001). On the other hand, PCR based microsatellites or simple sequence repeats (SSRs) are preferred for construction of genetic maps and molecular breeding in wheat due to their co-dominant inheritance, multi-allelic nature, high reproducibility and relative abundance (Somers, Isaac, and Edwards 2004). However, the genotyping of such type of markers is performed for individual markers, and are limited in addressing the kind of high-throughput required for genotyping large populations with large numbers of individual markers (Gupta *et al*., 2010).

To overcome such restrictions, hybridization-based methods using nucleic acids immobilised on solid-state surfaces have been developed. DNA chips were developed to analyse genotypes for single nucleotide polymorphisms (SNP) (Jaccoud *et al*., 2001). The SNP marker system is more valuable than SSR markers in building high-density genetic map, fine mapping of targeted genes, and gene cloning (Tian *et al.,* 2015) but the technology depends on intensive genomic sequencing and a high cost of analysis that cannot be matched in agriculture or basic research (Jaccoud *et al*., 2001). A technology with similar power but with a much greater breadth and much lower cost for DNA polymorphism analysis called Diversity Array Technology (DArT) was reported. Genetic marker analysis through Diversity Arrays offers a low-cost high-throughput, robust system with minimal DNA sample requirement capable of providing comprehensive genome coverage even in organism without any DNA sequence information (Jaccoud *et al*., 2001). Diversity Arrays Technology (DArT) can detect any type DNA variation at several hundred genomic loci in parallel without relying on sequence information (Wenzl *et al*., 2016) and has been applied successfully for polyploidy and complex genomes such as tetraploid and hexaploid wheat (Robbana *et al*., 2019).

2.7. **Overview of association mapping in plants**

The fundamental aim of genetics is to connect genotype to phenotype (Oraguze *et al*., 2007). There are two quantitative genetic approaches that are mainly used to identify associations between a genotype and phenotype namely linkage mapping and association mapping.

2.7.1. Linkage based QTL mapping

Linkage analysis is based on the co-segregation of marker alleles with phenotypic observations on families of a segregating population to identify quantitative trait loci (QTLs) that contain causal variants (Rafalski 2002). Linkage-based QTL mapping is a key approach for dissecting the genetic bases of complex agriculturally important traits in plants (Holland 2007) that uses bi-parental mapping populations to develop genetic linkage maps based on recombination during meiosis (Mammadov *et al*., 2012). In a linkage or bi-parental mapping population, the studied progeny are usually just a few generations away from the two common ancestors, which results in high linkage disequilibrium (LD), i.e. the non-random association of alleles between two loci (Rafalski 2002). Thus, even rather distant markers are found to co-segregate with the causal variant (Mauricio 2001; Maccaferri *et al*., 2008; MacCaferri *et al*., 2011; Mackay, Stone, and Ayroles 2009; Pasam *et al*. 2012). Linkage based QTL mapping is particularly useful in tagging QTL of rare variants and small-effect alleles. Although this approach has proven useful and remains an important tool in gene/QTL discovery (Mauricio, 2001; Maccaferri *et al*., 2008; MacCaferri *et al*., 2011; Mackay, Stone, and Ayroles 2009; Pasam *et al*., 2012; Korte and Farlow 2013), it has two fundamental limitations(Stich *et al*., 2006). One major limitation is that it can only capture restricted numbers of recombination events that occurred during the creation of the bi-parental mapping population. Efficient gene discovery with this method is largely limited to those loci that have large effects on quantitative trait variation. As a result, linkage maps have low mapping resolution inherent in samples with limited meiotic cross-over events, often in the range of 10-30 cM (Korte and Farlow 2013; (Oraguze *et al*., 2007). The identified QTL region can extend over several centiMorgans (cM) or more and contains hundreds of genes, making the identification of suitable candidates a very difficult task. Therefore, beneficial QTL alleles introgressed via marker-assisted selection are likely to suffer from linkage drag, i.e. the hitchhiking of deleterious loci with selected target loci (Oraguze *et al*., 2007). The second major limitation of the

traditional linkage based QTL mapping, the construction of mapping populations of recombinant inbred lines (RILs) through controlled crosses followed by several generations of selfing, is that only allelic diversity that segregates between the parents of the particular cross can be assayed throughout the longer, time consuming and costly research time scale which further limits the use of linkage mapping (Korte and Farlow 2013; Oraguze *et al*., 2007). The efficient gene discovery with this method is largely limited to those loci that have large effects on quantitative trait variation. To attain this aim, association mapping seeks to identify specific functional variants (i.e., loci, alleles) linked to phenotypic differences in a trait, to facilitate detection of trait causing DNA sequence polymorphisms and/or selection of genotypes that closely resemble the phenotype. The primary focus of association genetics is to identify polymorphism(s) that are located physically close to the causative trait polymorphism(s) (Oraguze *et al*., 2007).

2.7.2. Association mapping

Association or linkage disequilibrium (LD) based mapping is a powerful method and alternative methodology to the limitations inherent to linkage analysis in bi-parental mapping populations (Ross-Ibarra, Morrell, and Gaut 2007). Association mapping via linkage disequilibrium or LD (non-random association of alleles at different loci) offers promise in connecting genotype to phenotype. In association mapping, the natural variation that exists in the available germplasm can be utilized for mapping straightaway without a need to make crosses to generate segregating populations. Association genetics via LD mapping is an emerging field of genetic mapping that has the potential for resolution to the level of individual genes (alleles) underlying quantitative traits (Oraguze *et al*., 2007). Association Mapping has the potential to identify a single polymorphism within a gene that is responsible for phenotypic differences. In addition, many plant species have a high level of diversity for which association approaches are well suited to evaluate the numerous alleles available (Caliskan 2012; Flint-garcia, Thornsberry, and Iv 2003). It involves searching for genotype-phenotype correlations among unrelated individuals. Association Mapping has high resolution that is accounted for by the historical recombination accumulated in natural populations and collections of landraces, breeding materials and varieties. By exploiting broader genetic diversity, Association Mapping offers three main advantages over linkage mapping: mapping resolution, allele number and time saving in establishing a marker-trait association and its application in a breeding program (Caliskan 2012).

It is an approach that utilizes historical recombination and natural genetic variation within the ex situ conserved genetic resources; LD mapping identifies significant marker trait associations by exploiting linkage disequilibrium (LD) created through ancestral and evolutionary recombination events at the population level. LD-based association studies enable the most effective utilization of conserved genetic diversity among global germplasm resources (Abdurakhmonov and Abdukarimov 2008). It also provides the advantage of simultaneous evaluation of many alleles, hence, availability of broader genetic variation in diverse backgrounds for marker-trait correlations. Unlike linkage mapping, LD mapping leads to a higher resolution mapping because of using recombination events from a large number of meiosis that occurred throughout generations. Further, LD mapping enables exploitation of historical data for association. It is a timesaving and cost-effective approach compared to bi-parental populations (Abdurakhmonov and Abdukarimov 2008; Korte and Farlow 2013). In association mapping, the confounding effects of population structure and genetic relatedness present major limitations, if unaccounted for. This is because related individuals in a population share both causal and non-causal alleles and LD between non-causal sites can lead to false associations. Fortunately, these artefacts can be accounted for by using statistical models such as the mixed linear model (MLM) that can handle population structure by accounting for the phenotypic covariance that is due to population structure and genetic relatedness (Korte and Farlow 2013). Several studies showed the use of this approach to identify causal molecular polymorphism responsible for several traits. However, the association of SNPs with a trait still requires validation, as the SNP could be in disequilibrium with the causal factor, particularly if LD is high in the genomic region surrounding the gene. Thus, candidate gene approaches are generally utilized to eliminate putative candidates from more detailed functional studies (Flint-garcia, Thornsberry, and Iv 2003).

2.8. Approaches for association mapping

Two approaches are used for AM studies: whole-genome scans and candidate gene analysis (Thornsberry *et al*., 2001; Rafalski 2002;Kraakman *et al*., 2004; Rostoks *et al*., 2006; Pasam, Sharma, Malosetti, Eeuwijk, *et al*., 2012).

2.8.1. Genome-wide association mapping

Genome-wide association (GWA) mapping, also known as whole genome scanning approach, is a comprehensive approach in which genotyping is done for all possible genetic markers across the genome to systematically search the genome for causal genetic variation affecting the target trait(s) (Hirschhorn and Daly 2005). Genome-wide association mapping (GWAS) has a higher mapping resolution and higher potential to identify diagnostic markers than the traditional linkage mapping methods (Liu *et al*., 2017). The goal of GWAS is to discern genomic regions that could either be markers, genes or QTL associated with key morphological traits for marker-assisted breeding, gene discovery or gene introgression (Mwadzingeni *et al*., 2017). GWA identifies genomic regions throughout the genome associated with the trait of interest. A large number of markers are tested for association with various complex traits and prior information regarding candidate genes is not required. For many traits, such prior knowledge may not exist or may be very vague; thus, for these cases, genome wide AM is often used in association studies. The basis of genome-wide AM is to genotype a sufficient number of markers across the genome so the causal variants (i.e. the underlying gene(s) that control the trait) will likely be in LD with at least one marker. This approach is favoured in situations where LD extends for large distances, allowing for potential candidate regions associated with a trait of interest to be identified for further study (Remington *et al*., 2001). The extent of LD, therefore, is a critical factor in determining the number of markers needed to cover the genome and the mapping resolution that can be achieved. Association studies with high-density marker coverage; large sample size and minimum population structure offer great promise in complex trait dissection.

2.8.2. Candidate gene association mapping

A candidate gene is a coding sequence located in a chromosome region suspected of being involved in the expression of a trait whose protein product suggests that it could be the gene in question. Although the candidate gene approach has in some cases led to the identification of a few causal genes (Werner *et al*., 2005; Harjes *et al*., 2008; Zheng *et al*., 2008; Ramsay *et al*., 2011), it relies on some prior knowledge about the gene location and function. Candidate genes are selected based on prior knowledge of mutational analysis, biochemical pathway, or linkage and genome-wide association analysis of the trait of interest. However, this low-cost, hypothesis-driven, and trait specific approach will

inevitably miss the role of other unknown non-coding loci that may actually be the primary cause of the observed phenotypic variability (Zhu *et al*., 2008). The candidate gene approach directly tests the effects of genetic variants of a gene that may affect a particular trait. However, the candidate gene approach is limited by existing knowledge about the biology of the trait of interest and the genes underlying the QTL interval. This notwithstanding, the candidate gene approach is useful for quickly determining the association of a genetic variant with a trait and also for identifying genes of modest effect. Additionally, the increased availability of well-annotated genomes and the drastic reduction in sequencing costs are expected to facilitate the adoption of the candidate gene approach in future studies aimed to identify the loci governing the variability in traits of breeding interest.

CHAPTER III

3. Genetic Diversity of Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm from Ethiopia based on Phenotypic Traits

Abstract

Tetraploid wheat species from Ethiopia hold ample genetic variation, which could serve a base for improvement of wheat. A total of 196 Ethiopian tetraploid wheat (Triticum turgidum spp.) accessions, including 174 landraces and 22 improved cultivars, were evaluated at Sinana and Debrezeit during 2020 in simple lattice with two replications to assess the extent and pattern of genetic diversity based on 4 qualitative traits and 11 quantitative traits. Genetic diversity analysis using qualitative traits revealed the following phenotypic classes for seed color, spike density, awndness, and glume color: white, red, and purple; very lax, lax, intermediate, dense, and very dense; awnletted and conspicuous awns, and white, red to brown, and purple to black, respectively. All qualitative traits were highly significantly ($p<0.001$) variable across regions of collections and altitudinal classes; however, at the species level, only spike density was significantly ($p<0.05$) different. Intermediate diversity was observed in the entire genotypes of tetraploid wheat. The mean diversity index of the entire accessions, the regions of collections, altitudinal classes, and the species was 0.60±0.16, 0.57±0.07, 0.49±0.05, and 0.57±0.08, respectively. Shannon-Weaver diversity revealed high diversity in the entire landraces with respect to glum colour (0.78), spike density (0.61), and seed color (0.86); however, low diversity for awndness (0.15). All species exhibited high H' for glume color (0.77±0.01) and seed color (0.77±0.09); however, intermediate and low diversity for spike density (0.57±0.02) and awndness (0.17±0.03), respectively. In qualitiative traits, there was more diversity within than between at region level as well as altitudinal classes, except for awndness, in which between altitudinal was higher than within altitudinal variation. Combined analysis of variance using quantitative traits revealed highly significant variation ($p<0.001$) among genotypes and for genotype by location interaction for all traits except for the number effective tillers per plant. The observed mean and range values of the phenotypic traits revealed high variability in the accessions. Phenotypic Coefficient of Variation (PCV) and Genotypic Coefficient of Variation (GCV) values were high for grain yield, biomass yield, and harvest index. Grain yield showed highly significant ($p<0.001$) negative correlation with days to booting and days to maturity and positive correlation with all traits. The estimates of heritability (H^2) for grain yield and the number of spikelets per spike respectively ranged from 41.78 to 84.62 %. High values of genetic advance as a percentage of mean were recorded for the number of seeds per spikelet, the number of effective tillers per plant, grain yield, biomass yield, and harvest index. The number of seeds per spikelet gave a high value of genetic advance and heritability implying high genetic gain from its selection. The whole genotypes were grouped into seven c1usters, the number of genotypes ranging from 12 in the seventh cluster to 48 in the first cluster. The first and second principal components accounted for 19.74% and 15.96 of the total variation in the entire genotypes, respectively. The first five principal components with eigenvalues greater than one combined explained 61.21 % of the total variation in the entire genotypes. From the present result, it is concluded that there is a great variation in the tetraploid wheat germplasm of Ethiopia.

Keywords: Genetic diversity; Heritability; Landraces; Tetraploid wheat.

3.1. Introduction

Wheat is a commodity with a high market value that generates income for farmers in Ethiopia, the largest wheat producer in sub-Saharan Africa and a centre of diversity for the cultivated tetraploid wheat (Haile and Hammer 2013; Vavilov, 1929; Brasesco *et al*., 2019; Abate 2018). Despite this potential, Ethiopia remains a net importer of wheat due to the huge gap between production and consumption (Brasesco *et al*., 2019; Abate 2018) emanating from very low national yield (Haile *et al*., 2013) and increased demand for wheat (Zegeye *et al*., 2020). The mean productivity of wheat is 3 tons per hectare (tha^{-1}) (CSA 2021); however, it is below the global level (Alemu *et al*., 2019) and below the attainable yield of the crop that reaches up to 5t/ha (Nigus *et al*., 2022; Zegeye *et al*., 2020). Hence, low productivity continued to be the major challenge facing wheat (Alemu *et al*., 2019) as it could not cope with the ever-increasing demand. Accordingly, continued enhancement of wheat yield is crucially needed; which, in turn, necessitates the development of wheat varieties that are high yielding with the required quality and are stable under biotic and abiotic stresses. The development of new varieties requires a continuous supply of new germplasm as a source of desirable genes (Meseret *et al*., 2020) for which tetraploid wheat landraces are the potential source.

Tetraploid wheat of Ethiopia holds huge variation, containing useful genes valuable for future plant breeding work (Jack R. Harlan 1969). The pattern of variation provides substantial information and is basic to an understanding of the germplasm of a crop. The variation, within and among populations, must be studied in detail and depth (Harlan 1971) and properly documented for the final success of the conservation program. The importance of studying patterns and distribution of variation in tetraploid wheat germplasms was emphasized by (Belay *et al*., 1993, 1997; Hailu *et al*., 2010; Tesemma *et al*., 1993; Bechere *et al*., 1996; Bekele 1984; Ayana and Bekele 1999) in their studies. More than 7000 landraces of tetraploid wheat were collected and conserved by the Ethiopian Biodiversity Institute; however, only limited portions of them were studied (Negisho *et al*. 2021). Detailed knowledge about the collections is needed for the efficient utilization of the genetic potential contained in the germplasm collections (Beuselinck and Steiner, 1992; Ayana and Bekele 1999). Therefore, assessing the genetic diversity of tetraploid wheat landraces is crucial to address the different breeding purposes such as stable and high grain yield, grain quality and rust resistance.

Genetic diversity is a critical input for successful adaptation and sustained improvement in crops (Marone *et al*., 2012) including tetraploid wheat and is the basis for genetic improvement (Huang *et al*., 2002) and provides opportunity to develop new and improved cultivars with desirable characteristics (Govindaraj *et al*., 2015). Assessing the genetic diversity of tetraploid wheat landraces is a prerequisite for plant breeding and genetic resource conservation programmes (Laidò *et al*., 2013) and crucial to address the different breeding purposes such as stable and high grain yield, grain quality, and rust resistance. Knowledge of genetic diversity is very important to identify, conserve, and efficiently utilize useful materials from the existing landraces (Alamerew *et al*., 2004; Ruiz *et al*., 2012), is the basis for parents selection in crossing program, and has significant impact on improvement of crops (Zhang et al., 2011). Information on available genetic resources, their geographical locations, and understanding of their relationships can be used to gain insight into population divergence (Mengistu *et al*., 2016); however, the available information is not exhaustive and representative (Meseret *et al*.,2020). Such information is derived from phenotypic as well as genotypic studies using either morphological or molecular markers or a combination of both for rigorous characterization of genotypes (Kidane *et al*., 2017). A better knowledge of the collected material can be obtained using morphological, biochemical, and molecular markers (Teklu and Hammer 2008), which can help us in the assessment of genetic variability among genotypes within landraces, populations, and species (Eticha *et al*., 2005).

Morphological markers have been in use for diversity analysis starting in the era of selection-based plant breeding and continue to play a role even today when molecular markers are being used. They are the alternative to their molecular counterparts for assessing genetic diversity. Highly heritable morphological characters are still applicable for germplasm characterization and management and can be analysed anywhere in the world without needing laboratory capacity, or the expense of sophisticated equipment and chemicals (Hailu 2011). To unlock genetic diversity of tetraploid wheat of Ethiopia attempts were made by many researchers using morphological diversity (Getachew and Worede 1991; Eticha *et al*., 2005; Hailu *et al*., 2006; Belay *et al*., 1993; Bechere *et al*., 1997; Teklu and Hammer, 2008; Getachew et al. 2017; Mengistu *et al*., 2015). Including durum wheat, more than 7000 tetraploid wheat landraces are collected and maintained at EBI but only limited portion of these collections are studied (Negisho *et al*., 2021). Therefore, as assessing the phenotypic diversity of Ethiopian tetraploid wheat is crucially

needed, this study was initiated to study the extent and patterns of genetic diversity based on phenotypic traits.

3.2. Materials and Methods

3.2.1. Description of the study areas

Genetic diversity analysis for phenotypic traits was conducted at Debrezeit and Sinana Agricultural Research Centers. Debrezeit Agricultural Research Center is located at an altitude of approximately 1,900 m above sea level, with latitude of 80 44' N and longitude of 380 85' E. Sinana is located at an altitude of 2400 m.a.s.l. Sinana has a range of mean annual rainfall of 563-1018mm with minimum and maximum temperature of 7.9°C and 24.3°C, respectively.

3.2.2. Planting materials and experimental procedures

A total of 196 tetraploid wheat (*Triticum turgidum* spp.) genotypes, representing 174 landraces and 22 varieties, collected from different parts of the country (Figure 3.1) were used in this study (Annex 1). The landraces used were originated from the different Ethiopian wheat producing regions: Shewa, Jima, Bale, Tigray, Wello, Gonder, Gojam, Agaw Awi (Figure 2.1). The released varieties and 40 of the landraces were obtained from Debrezeit Agricultural Research Centre. The remaining 134 landraces were obtained from Sinana Agricultural Research Centre. The landraces were collected by the Biodiversity Institute of Ethiopia (IBE). Ten plants were randomly selected and tagged for phenotypic data collection. Highly heritable morphological traits that are mostly used in the breeding of wheat were selected from the descriptors for wheat (IBPGR 1985) for phenotypic characterization of the tetraploid wheat. Genotypes were grown in two rows of 1 meter length and 20 cm between-row spacing in a simple lattice design with two replications. Agronomic recommendations used: 100 kg urea and 150 kg DAP per hectare and three times hand weeding was applied. For data analysis, landraces from regions with small sample sizes were included into adjacent regions to reduce experimental errors due to small sample sizes. In doing so, landrace from Jimma along with, one landrace with unknown location, were included into Shewa, five landraces from Gonder and two from Agaw Awi were included into Gojam, two landraces from Tigray into Wello and three landraces from Bale were included into Arsi.

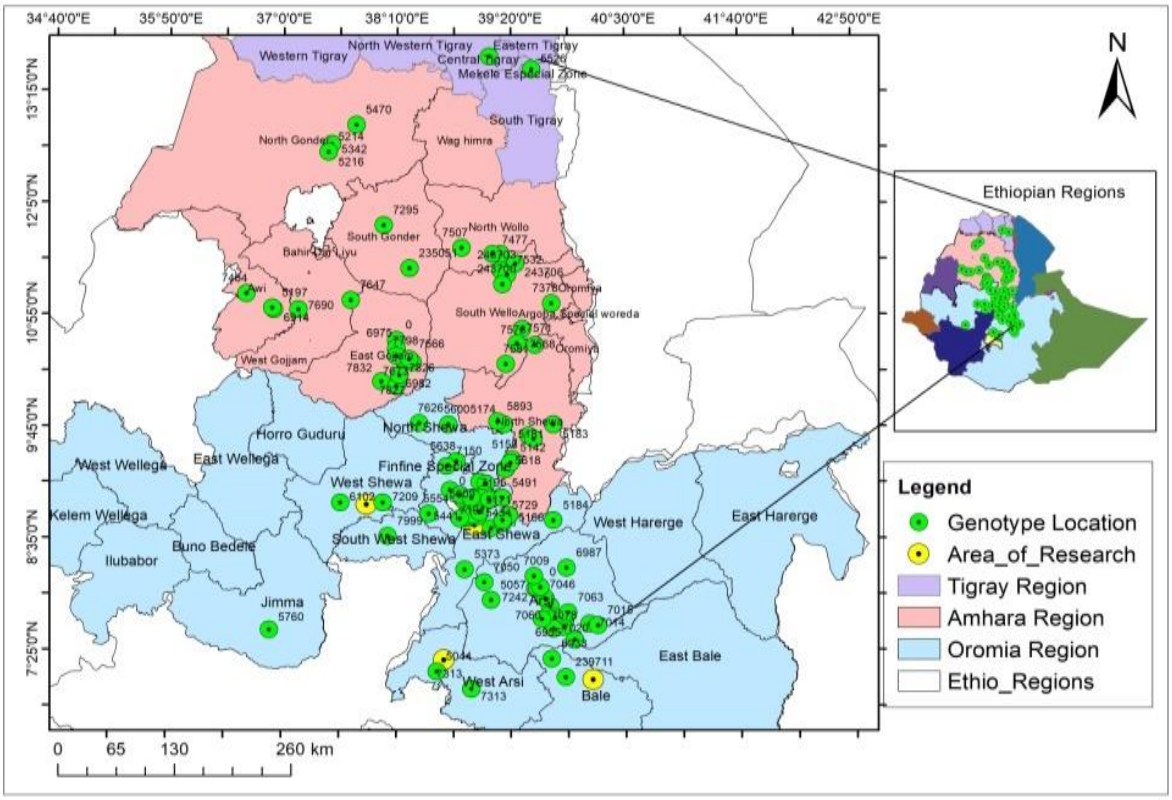


Figure 3. 1 Geographical map of Ethiopia, indicating areas of collection of the landraces

3.2.3. Qualitative morphological traits data collection and analysis

The genotypes were assessed for four qualitative morphological traits: spike density, awndness, glume color, and seed color based on descriptors for wheat (IBPGR 1985) to compute Shannon-Weaver diversity as described in Table 3.1.

Spike density: a visual measured of the density of a spike measured on a 1-9 scale: 1 (very lax), 3(lax), 5 (intermediate), 7 (dense) and 9 (very dense).

Awndness: presence or absence of awns measured on a 0-7 scale: 0 (awnless), 3 (awnletted or short awns), and 7 (awned or conspicuous awns).

Glume color: observed on the outer glume: 1 (white), 2 (red to brown), and 3 (purple to black).

Seed color: visual observation of seed color: 1 (white), 2 (red), and 3 (purple).

Table 3. 1 Characters used for the study, their codes, and descriptions

Morphological traits	Categories	
Qualitative traits	Code	description
Spike density	1	very lax
	3	lax
	5	intermediate
	7	dense
	9	very dense
Seed color	1	white
	2	red
	3	purple
Glume color	1	white
	2	red to brown
	3	purple to black
Awndness	0	Awnless
	3	Awnletted/short awns
	7	awned/conspicuous awns

Frequencies of the phenotypic classes of each character within the entire genotypes, altitudinal classes, species level, and regions of collections were calculated in excel. Hierarchal clustering of accessions was performed based on regions of collections and altitudinal classes using R software for the standardized data to mean zero and unity variance. The amount of genetic variation was determined using the Shannon-Weaver diversity index as suggested by Jain *et al*. (1975) using the formula: $H' = -\sum_{i=1}^{n} PixLnPi$ where n = the number of phenotypic classes for a character and P_i = the genotypic frequency as the percentage proportion of the total entries in the i^{th} class. The diversity index was estimated for each trait at specific accession level, regional level, and altitudinal level. H' values were standardized using the formula: $H' = \frac{H}{Hmax}$ to express the values of H' in the range of 0–1 and were used in the analysis of variance of diversity-based regions of collections, species, and classes of altitudinal range. The diversity index was categorized into high ($H' > 0.60$), intermediate ($0.40 \leq H' \leq 0.60$), or low ($0.10 \leq H' <40$) based on Eticha *et al*. (2005) illustration.

3.2.4. Quantitative morphological traits data collection

The 196 genotypes (appendix 1 and 2) were assessed for 11 quantitative morphological traits following descriptors for wheat (IBPGR 1985). Days to booting, days to maturity, seed yield, and biomass yield were recorded on a plot basis.

Plant height: Height of plant at maturity, measured in cm from ground to top of spike.

Spike length: the length of spike from the base of spike to the tip of spike measured in cm.
Days to booting: counted as days from sowing to 50% of plants in booting stage.

Days to maturity: counted as days from sowing to 50% of plants at physiological maturity

Number of spikelets per spike: The average number of spikelets per spike from five typical spikes randomly selected from a growing accession.

Number of seeds per spikelet: The average number of seeds from a spikelet - obtained from the central portion of the five randomly selected typical spikes.

Number of effective tillers per plant: the number of tillers bearing spike from the five randomly selected plants.

Biomass yield per plot: dry weight of the above ground wheat per plots taken at harvest.

Grain yield per plot: the grain weight of all plants grown per plot taken from each genotype, moisture content adjusted to 12.5%.

Thousand kernel weight: the weight of thousand grains of wheat genotypes.

Harvest index: the ratio of dried grain weight per plot divided by above ground biomass at 12% moisture content.

3.2.5. Quantitative data analysis

Analysis of variance (ANOVA) for each location was carried out using the PBIB test in R by considering genotypes and block as fixed and random factors, respectively. In each location, the observed phenotypic response of the i^{th} genotype in the j^{th} replication and l^{th} sub-block was computed using the following model:

$$Y_{ijl} = \mu + g_i + y_j + bl(j) + \varepsilon_{ij}$$

where y_{ijl} = the observed phenotype, μ = the grand mean, g_i = fixed effect of the i^{th} genotype, y_j = effect of the j^{th} replication, $bl_{(j)}$ = random effect of the l^{th} block nested within the j^{th} replication, and ε_{ijl} = random error term.

The combined ANOVA across the locations was executed by considering genotype as a fixed effect and the block and location as random effects according to the following model:

$$Y_{ijkl} = \mu + g_l + r_{ijk} + l_i + b_{ijkl} + (gl)_{il} + \varepsilon_{ijkl}$$

where Y_{ijkl} = observed response of genotype l and replication j of block k of location i; μ = grand mean; g_l = fixed effect of genotype l; r_{ijk} = effect of replication j in location i; l_i= random effect of location i that is ~NID(0, δ^2_e); b_{ijkl} = random effect of block k nested within replication j in location i that is ~NID(0, δ^2_b); $(gl)_{il}$ = random effect of the interaction between genotype l and location i that is ~NID(0, δ^2_{gl}) + ε_{ijkl} = random residual effect that is ~NID(0, δ^2_ε).

Homogeneity of the error mean square (MS) from individual analysis of variance at Sinana and Debrezeit was checked following the F-max technique of Hartley (1950), which is based on the ratio of the larger mean square error (MSE) from the separate analysis of variance to the smaller MSE. The error variance is declared as homogenous if the larger MSE is not three times greater than the smaller MSE (Gomez and Gomez, 1984). After deciding that the error variances were homogeneous for specific triat following the formula of Gomez and Gomez (1984) described as: F=Larger error MS/smaller error MS, the combined analysis of variance was performed. The statistical signficance between genotypes is decided based on p-value that corresponds to the F stastics. If p-value is less than the specified alpha (α) level, the null hypothesis is rejected and the difference between the genotypes concluded significant; however, if the p-value is not less than the specified alpha (α) level, the null hypothesis is accepted and we conclude that the genotypes showed statistically no significant difference (https://statisticalpoint.com/anova-f-value-p-value).

Phenotypic, genotypic, environmental, and genotype by environment interaction coefficient of variation, broad sense heritability (H^2), and genetic advance were the parameters assessed.

The variability of each quantitative trait was estimated by simple statistical measures such as mean, range, phenotypic and genotypic variances, and coefficient of variation. The phenotypic and genotypic variation and coefficient of variations were calculated following the formula suggested by Singh and Chaudhary (1985) and Allard (1960) as follows:

Genotypic variance (δ^2_g)

$$\delta g^2 = \frac{MSg - MSgl}{rl}$$

where MS_g = mean square of genotype, MS_{gl} = mean square due to genotype by environment interaction, l = number of locations and r = number of replications.

Environmental variance ($\delta^2 e$) = MSe where MSe= combined error mean square.

Genotype by environment interaction variance (δ^2_{gl})

$$\delta gl^2 = (MSgl - MSe)/r$$

where MS_{gl} = mean square due to genotype by environment interaction, and MSe = combined error mean square.

Phenotypic variance (δ^2_p) $\delta p^2 = \delta g^2 + (\delta gl^2/l) + (\delta e^2/rl)$

Estimates of coefficient of variation were obtained as follows.

Phenotypic coefficient of variation (PCV)

$$\boldsymbol{PCV} = \frac{\sqrt{\delta p^2}}{\mu}\boldsymbol{x100}$$

where PCV = phenotypic coefficient of variation, δ^2_p = phenotypic variance and μ= population mean for the trait considered.

Genotypic coefficient of variation (GCV)

$$\boldsymbol{GCV} = \frac{\sqrt{\delta g^2}}{\mu}\boldsymbol{x100}$$

where GCV = genotypic coefficient of variation, δ^2_g = genotypic variance and μ = population mean for the trait considered.

Environmental coefficient of variations (ECV)

$$\boldsymbol{ECV} = \frac{\sqrt{\delta e^2}}{\mu}\boldsymbol{x100}$$

Genotype by environment interaction coefficient of variation (GECV)

$$\boldsymbol{GECV} = \frac{\sqrt{\delta gl^2}}{\mu}\boldsymbol{x100}$$

where, δ^2_{gl} = genotype by environment interaction variance and $\boldsymbol{\mu}$= population mean for the trait considered.

Broad sense heritability (H^2) and genetic advance

Heritability in broad sense, for the two locations, was estimated based on the formula given by (Allard 1960).

$H^2 = \left(\frac{\delta 2g}{\delta 2p}\right) X\ 10$, where $\delta 2p = \delta 2g + \left(\frac{\delta 2gl}{l}\right) + \left(\frac{\delta 2e}{rl}\right)$, $\delta^2 e$ = error variance, l= the number of locations and r= the number of replications.

Expected genetic advance under selection, assuming a selection intensity of 5%, was computed following the formula developed by Allard (1960) as: $GA = (K)\ (\delta p)\ (H^2)$, where GA = expected genetic advance, K = selection differential that varies depending on the selection intensity and stands at 2.056 for selecting 5% of the genotypes, δp = phenotypic standard deviation, and H^2 = heritability (in a broad sense).

Genetic advance as percentage of the mean was calculated as $GA\ (\%\ of\ mean) = \left(\frac{GA}{\mu}\right) x100\%$, where GA = genetic advance and μ= population mean for the trait considered.

Cluster analysis: Hierarchal clustering for the entire data set, regions of collections, altitudinal classes, and at species level was performed based on ward.D linkage method with Euclidian distance, obtained by dist () function, using R software version 4.1.1(Harris 2018). The distances between clusters and within clusters were calculated using the average linkage method of Euclidean distance based on their centroids.

Principal component analysis: Principal component analysis was performed for the entire data set, regions of collections, altitudinal classes, and at species level to identify the most important traits contributing to the total variations observed among the accessions, regions of collections and altitude classes using princomp () function in R software version 4.1.1 (Harris 2018)**.** As suggested by Johnson *et al.*(2007), principal components with Eigen values greater than one were considered.

Correlation coefficient analysis: Pearson correlation analysis of quantitative traits was performed for quantitative traits using R software (version 4.1.1) (Harris 2018)**.**

Comparison of selected genotypes with the original population: to compare with the original population, the means of the top 5% genotypes for each trait were independently computed. The student t-test table was used to determine the significance of the difference between the sample mean and population parameter. When the calculated t-value is higher

than the tabulated t-value, the difference is deemed significant. To compare the performance of the 5% best-selected genotypes with the size of a population, the absolute t-value was obtained using the student's t-test formula as follows:

$T = (\bar{x} - \mu)/(\frac{S}{\sqrt{n}})$ Where n is the number of genotypes selected from the size of a population for better performance, $\bar{x}$ is the mean of the genotypes that were selected, S is the sample standard deviation, and μ is the mean of the size of the population.

3.3. Results and Discussion

A total of 196 genotypes of tetraploid wheat (*Triticum turgidum* spp), including 174 landraces along with 22 improved cultivars, of Ethiopia (Annex 1) were evaluated to assess the extent and pattern of genetic diversity across regions of collections, species level, and altitudinal classes using diversity index, genetic variation, genetic advance, heritability, correlation analysis, principal component analysis, and cluster analysis on the basis of phenotypic traits using four qualitative and eleven quantitative traits at Sinana and Debreziet in simple lattice with two replications during 2020.

3.3.1. Qualitative traits diversity

3.3.1.1. Analysis of variance and diversity index estimation of the qualitative traits

Mean squares and Shannon-Weaver diversity index estimation of four qualitative traits is given in Table 3.2. The ANOVA revealed that all qualitative traits were highly significantly ($p<0.001$) variable across regions of collections and altitudinal classes; however, at the species level, glume color, awndness, and seed color were not significantly different while spike density was significantly ($p<0.05$) different (Table 3.2). The estimated mean diversity for the entire genotypes was 0.60±0.16. In line with this study, Jain *et al*. (1975) and Negassa (1986) reported high mean diversity (0. 87 ± 0. 04, 0.81) in tetraploid wheat Ethiopia and in wheat of Ethiopia, respectively. The estimated diversity index (H') for the whole accessions was 0.15, 0.78, 0.61, and 0.86 for awndness, glume color, spike density, and seed color, respectively. Incontrast, Bechere *et al*.(1996) reported low mean diversity for spike density (H' = 0.24).

The mean diversity index was 0.60±0.15, 0.59±0.19, 0.51±0.14, and 0.60±0.16 respectively for Shewa, Gojam, Arsi, and Wello. In line with this study, Belay *et al*. (1997) and Negassa (1986) respectively reported similar result on mean diversity for Shewa collections in tetraploid wheat and for Gojam (0.73) and Arsi (0.69) from wheat of Ethiopia. The mean diversity index across the regions of collections, altitudinal classes, and the species level were respectively 0.57±0.07, 0.49±0.05, and 0.57±0.08. The overall mean diversity for the whole qualitative traits ranged from 0.51±0.14 to 0.60±0.16, 0.22±0.27 to 0.60±0.32, and 0.54±0.11 to 0.59±0.17 at regional, altitudinal, and species level, respectively. An intermediate mean diversity index for the traits was obtained at altitudinal class 2021-2200, 2201-2380, 2381-2560, 2561-2740, 2741-2920, and ≥2921; however, low mean diversity was observed at altitudinal class ≤1840 and 1841-2020. In

line with this study, Hailu *et al*. (2010) reported the low diversity index over all the traits for accessions obtained below 2000 m.a.s.l.

The mean diversity index across regions of collections indicated that seed color (0.84±0.05) and glume color (0.77±0.01) showed high diversity; however, spike density (0.54±0.04) and awndness (0.14±0.02) had intermediate and low diversity, respectively. In line with the present study, Hailu *et al*. (2010) reported similar result on spike density, glume color, and seed color.

The estimated mean diversity index for glume color, spike density, awndness, and seed color respectively was 0.75, 0.63, 0.16, and 0.85 at Shewa; 0.77, 0.54, 0.09, and 0.95 at Gojam; 0.75, 0.43, 0.14, and 0.70 at Arsi, and 0.80, 0.54, 0.17, and 0.87 at Wello. In line with the present study, Hailu *et al*. (2010) reported high values of H' for spike density for the landraces from Bale region, for glume color for landraces from Wello, and for seed color for landraces from Shewa.

Altitudinal class mean diversity index was low for awndness (0.07±0.03); intermediate for spike density (0.53±0.04) and high for seed color (0.74±0.05) and glume color (0.62±0.11). In line with this study, Hailu *et al*. (2010) reported high H' value for spike density. At the species level, high mean diversity index (H') was recorded for glume color (0.77±0.01) and seed color (0.77±0.09); however, intermediate and low mean diversity was observed for spike density (0.57±0.02) and awndness (0.17±0.03), respectively. Glume color, spike density, and awndness, respectively, showed high, intermediate, and low mean diversity across species; however, seed color showed high mean diversity for *Triticum durum* (0.88) and *Triticum turgidum* (0.82) and intermediate for *Triticum aethiopicum* (0.60).

Table 3. 2 Mean squares and Estimates of the Shannon-Weaver diversity Index (H') for phenotypic traits in 174 landraces based on altitudinal class, regions of collections, and species

		Traits				
		GC	SD	Aw	SC	Mean
Population level		0.78	0.61	0.15	0.86	0.60±0.16
Region	Shewa	0.75	0.63	0.16	0.85	0.60±0.15
	Gojam	0.77	0.54	0.09	0.95	0.59±0.19
	Arsi	0.75	0.43	0.14	0.70	0.51±0.14
	Wello	0.80	0.54	0.17	0.87	0.60±0.16
Mean		0.77±0.01	0.54±0.04	0.14±0.02	0.84±0.05	0.57±0.07
MS (df=3)		0.53***	1.89***	1.60***	1.03***	
Altitudinal class	≤1840	0.29	0.57	0	0.58	0.36±0.28
	1841-2020	0	0.29	0	0.57	0.22±0.27
	2021-2200	0.77	0.55	0	0.62	0.49±0.34
	2201-2380	0.81	0.61	0.20	0.73	0.59±0.27
	2381-2560	0.77	0.57	0.18	0.89	0.60±0.31
	2561-2740	0.76	0.59	0.16	0.90	0.60±0.32
	2741-2920	0.73	0.61	0	0.90	0.56±0.39
	≥2921	0.83	0.47	0	0.69	0.50±0.36
Mean		0.62±0.11	0.53±0.04	0.07±0.03	0.74±0.05	0.49±0.05
MS (df=7)		0.016***	0.044***	0.06***	0.028***	
Species	TE	0.76	0.56	0.24	0.60	0.54±0.11
	DU	0.79	0.54	0.13	0.88	0.59±0.17
	TR	0.77	0.60	0.15	0.82	0.59±0.15
Mean		0.77±0.01	0.57±0.02	0.17±0.03	0.77±0.09	0.57±0.08
MS (df=3)		0.007	0.03*	0.001	0.007	

Remarks: TE= *Triticum aethiopicum*, DU= durum, and TR= *Triticum turgidum*, MS= Mean Square, GC= Glum color, SD= spike density, SC= seed color, and Aw= awndness.

3.3.1.2. Regional, species level, and altitudinal distribution of qualitative traits

The percentage frequency distribution of the qualitative traits is given in Annex 2. In the present study, the phenotypic classes observed for seed color, spike density, awndness, and glume color were respectively white, red, and purple; very lax, lax, intermediate, dense, and very dense; awnletted and conspicuous or large awns, and white, red to brown, and purple to black (Figure 3.2 and Annex 2).

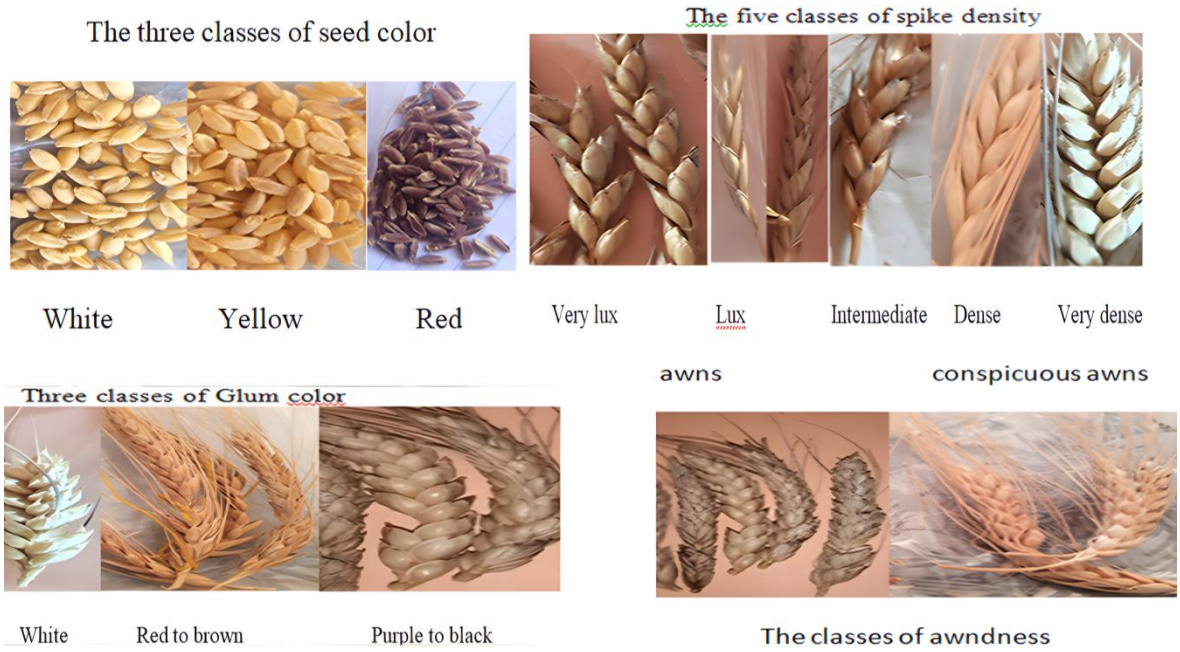


Figure 3. 2 Major phenotypic classes of qualitative traits recorded in tetraploid wheat landraces and improved cultivars of Ethiopia.

The frequency distribution of glume color showed that the white color was dominant across species, altitudinal classes and in the growing regions except in Arsi, where the red to brown color type was dominant. The purple to black glume color was not observed in the ≤1840 and 1841-2020 altitudinal classes. In line with this study, Bekele (1984) reported that the white glume color is dominant in tetraploid wheat of Ethiopia and Pecetti *et al*. (1992) reported that the white glume color dominates in durum wheat of Ethiopia. In line with the present study, Belay *et al*. (1997) reported that black glume color was less frequent; however, their report disagrees with this study in that white seed color was dominant only in Tigray collections and less present in other regions of collections such as Shewa and Gojam, In line with this study, Jain *et al*. (1975) reported that white glume color predominates in world collections of durum wheat and black color is less frequent.

In the case of spike density, the intermediate type was dominant in *durum* and *turgidum* species, while lux type was dominant in *aethiopicum* species. Very lux, intermediate, and dense types were equally found at altitudinal class ≤1840. Very dense type was dominant at altitudinal class 1841-2020. Intermediate type was dominant at altitudinal classes 2021-2200, 2201-2380, 2561-2740, and 2741-2920. Dense type was dominant at altitudinal classes 2381-2560 and ≥2921. Intermediate type of spike density was dominant in Arsi and Gojam collections while the dense and intermediate types were equally observed in Shewa and Wello collections. The very lax and the lax types were rarely observed across regions

of collections, altitudinal classes, and species. Opposite to the present study, the report of Belay et al., (1997) indicated that dense spike was predominant in all regions. Conspicuous awns were dominant in all the regions of collections, altitudinal classes, and across the species category in the awndness types. On the other hand, the average frequency distribution of seed color indicated that the red type was dominant followed by the purple type across the regions of collections, altitudinal classes, and species level, except in Wello where purple type was dominant and in altitudinal class 2 (1841-2020) at which the three seed color types are equally found. The purple type or yellow color is most preferred for pasta making and can be used in screening experiment for pasta production (Petrova 2007).

The average frequency distribution of seed color indicated that across regions of collections, altitudinal classes, and species, the red type was dominant followed by the purple type except for Wello collections where purple type was dominant. Despite the current preference for the white and purple types in the international market, the landraces in the present study were dominated with the red type followed by the purple type seed color indicating that the farmers maintained these landraces for a long period for their local utilization. In line with this study, Jain *et al*. (1975) reported that red types seed are dominant in tetraploid wheat of Ethiopia. Shannon-Weaver diversity index of seed color for the entire genotypes (0.86), across regions collections (0.84±0.05), altitudinal classes (0.74±0.05), and species level (0.77±0.09) revealed that there was high diversity in seed color among the tetraploid wheat genotypes of Ethiopia. In line with this study, Ayana (2015) reported high diversity index for seed color (0.74) in Ethiopian tetraploid wheat. Opposite to this study, Ayana (2015) reported white seed color is dominant in tetraploid wheat of Ethiopia. In line with this study, Belay *et al*. (1997) reported that white seed color is less frequent in Shewa and Gojam; however, their report disagrees with the present study that purple type is dominant in Shewa.

While considering the awndness class, conspicuous awns were dominant in all the regions of collections, altitudinal classes, and across species categories. Awnless types were not observed across the regions of collections, altitudinal classes, and species level.

3.3.1.3. Clinal variation of qualitative traits

Clinal variation of qualitative traits is given in Figure 3.3. Glume color, spike density, and seed color showed a decreasing trend from altitudinal class 1 to altitudinal class 2; however, awndness remains the same until altitudinal class 3. Seed color showed an

increasing trend in its mean Shannon-Weaver diversity starting from altitudinal class 2 until altitudinal class 7, after which it decreased. Glume color showed an increasing trend in its Shannon-Weaver diversity from altitudinal class 2 until altitudinal class 4, after which it showed a decreasing trend until altitudinal class 7 and then, increased in altitudinal class 8. Spike density showed an increasing trend from altitudinal class 2 until altitudinal class 4 and decreased in altitudinal class 5 and started increasing from altitudinal class 6 until class 7, after which it decreased. Awndness remained monomorphic in its mean Shannon-Weaver diversity from altitudinal class 1 to class 3 and it increased in altitudinal class 4, after which it showed a decreasing trend until class 6 and remained monomorphic until class 8.

3.3.1.5. Between and within region(s) and altitudinal classes' diversity: The partitioning of the diversity index (H') into within and between the regions of collections and altitudinal classes was executed to study the sources of variation in the entire populations (Table 3.3).

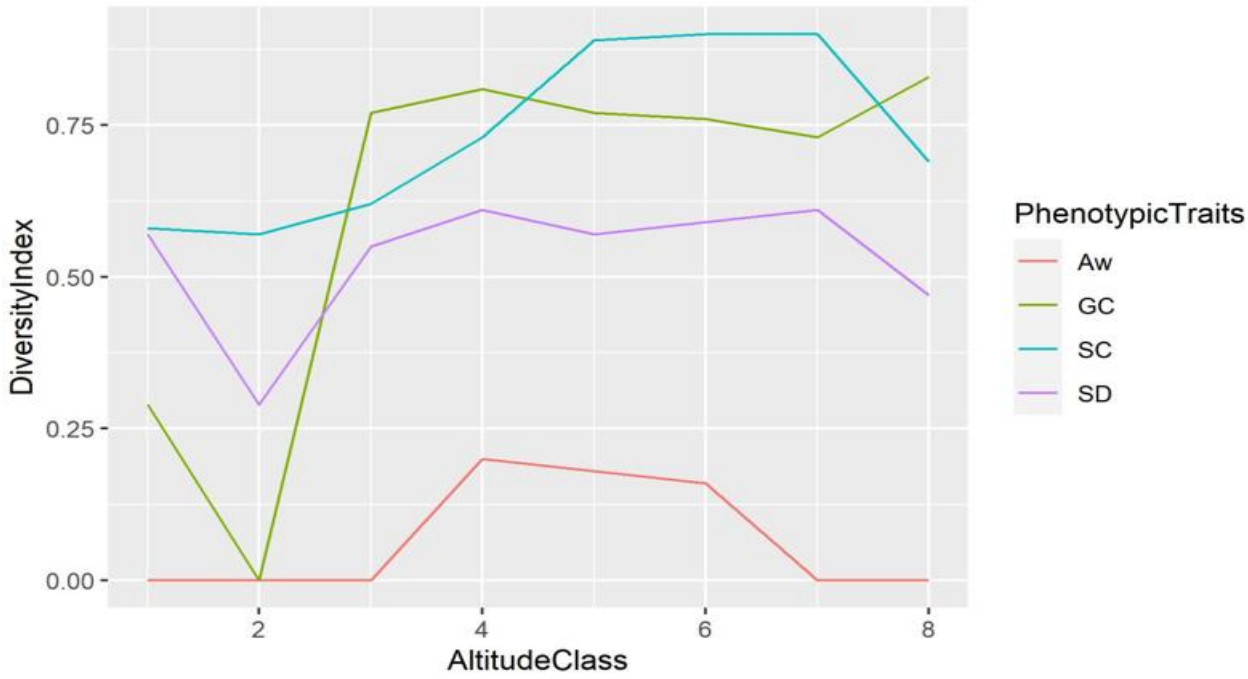


Figure 3. 3 Clinal variation in qualitative traits of Ethiopian tetraploid wheat germplasm (AW= Awndness, GC= Glume Color, SC= Seed Color, and SD= Spike Density).

Pooled over all traits, within region diversity and between regions diversity were 95% and 5%, respectively implying that there are enough variations within the region as compared to between regions for seed color, glume color, awndness, and spike density. Similarly,

pooled over traits, within altitudinal class diversity accounted for 75% of the total variation; however, only 25% of the total variation was accounted for by between altitudinal classes. There was more variation within altitudinal classes than between altitudinal classes for seed color, glume color, and spike density; however, within altitudinal diversity was less than between altitudinal diversity for awndness. In line with the present study, Negash (2015) reported that within region diversity being larger than between region diversity and within altitude variation is larger than between altitudinal variation in tetraploid wheat population of Ethiopia. In line with the present study, Teklu and Hammer (2008) reported that variation is attributed to within region variation. Opposite to the present study, Hailu *et al*. (2010) reported that the variation is due to among accessions rather than within accessions.

In the present study, a pattern of phenotypic diversity was conducted on 174 tetraploid wheat landraces along with 22 improved cultivars using the Shannon-Weaver diversity index across regions of collections, altitudinal levels, and species level. The overall mean diversity (H') recorded for the entire genotypes was intermediate (0.60±0.16). Owing to the sensitivity of Shannon-Weaver diversity index comparison of different studies with different descriptors having different classes needs care (Grenier *et al*., 2004; Teklu and Hammer, 2008). In contrast to this study, Jain *et al*.(1975) in a world durum wheat collection, Bechere *et al*. (1996), Hailu *et al*., (2010), Negassa (1986) and Negash (2015) on tetraploid wheat of Ethiopia, and Negassa (1986) on durum wheat of Ethiopia reported high mean diversity; however, the number and type of traits used for characterization were different.

Table 3. 3 Estimates of the Shannon-Weaver diversity index (H') according to the within and between tetraploid wheat regions of collections and altitudinal classes for the qualitative traits assessed on 174 landraces

		Traits				
Diversity index		GC	SD	SC	AW	Mean
H'		0.78	0.61	0.86	0.15	0.60±0.16
Region	Hcr	0.77	0.54	0.84	0.14	0.57±0.16
	H'_{cr}/H' (within)	0.99	0.89	0.98	0.93	0.95±0.02
	$(H'-H_{cr})/H'$(between)	0.01	0.11	0.02	0.07	0.05±0.02
Altitude	Hca	0.62	0.53	0.74	0.07	0.49±0.15
	H'ca/H' (within)	0.79	0.87	0.86	0.47	0.75±0.10
	$(H'-H_{ca})/H'$ (between)	0.21	0.13	0.14	0.53	0.25±0.10

H' = diversity index for each trait calculated from the entire dataset; H_{rc} and H_{ca} = average (arithmetical) diversity index of each trait for the regions of collections and altitudinal

classes respectively; Hcr/H' and Hca/H' = proportion of diversity within regions and altitudes respectively and (H' - Hcr)/H' and (H' - Hca)/H' = proportion of diversity between regions of collections and altitudinal classes in relation to the total variation.

Three classes of seed color (white, purple/yellow, and red) were observed implying that the different seed-colored genotypes could be utilized for different end use purposes (Hailu *et al*. 2010). The different phenotypic classes of landraces were maintained by farmers for a long period of time owing to their different end use qualities (Geleta and Grausgruber, 2014; Ayana, 2015) and grain color is one of the grain quality traits influencing wheat yield and quality (Alemu *et al*., 2020) and finally deciding the conservation and utilization of the specific crop. Based on the end-product, different color wheats may be desirable in different markets (Tipples *et al*., 1994) making all the different colored wheat useful at both farmers and industry level. Thus, conservation and utilization of the genotypes for different breeding objectives are recommended.

3.3.2. Quantitative traits analysis

3.3.2.1. Combined analysis of variance

A total of 196 tetraploid wheat (*Triticum turgidum* spp.), including 174 landraces and 22 varieties, were assessed for phenotypic diversity based on eleven quantitative traits at Debrezeit and Sinana Agricultural Research centres during 2020. After checking for homogeneity of the error MS from individual analysis of variance at Sinana and Debrezeit during 2020 following the formula of Gomez and Gomez (1984) described as: the ratio of the larger mean square of error (MSe) to the smaller MSe based on the F-max technique of Hartley (1950), the error variances were homogeneous for all the traits studied. Moreover, the genotypes were assessed for the principal component analysis and cluster analysis based on the fifteen phenotypic traits (4 qualitative and 11 quantitative traits). Table 3.4 illustrates the variance results from a pooled analysis of eleven phenotypic traits for 196 genotypes of tetraploid wheat genotypes at Sinana and Debrezeit during 2020.

Analysis of variance of quantitative traits revealed highly significant variation ($p<0.001$) for all traits among genotypes and genotype by location interaction except for the number of effective tillers per plant. All traits showed highly significant ($p<0.001$) variation across locations. The highly significant variation of trait across locations and among genotype by location interaction emphasized the role of locations and environment in shaping the traits (Studnicki *et al*., 2016). The highly significant ($p<0.001$) variation among the genotypes

for all traits other than the number of effective tillers per plant is an indication of the presence of genetic variation among the genotypes which in turn suggests that selection of lines can be effective in improving both yield and quality traits (Azene *et al*., 2020). In line with this study, Azene *et al.*(2020) reported signficant variation among genotypes of durum wheat in Ethiopia. The signficant ($p<0.001$) variation acorss locations and genotype by location interaction suggest that the significant phenotypic variation among the tetraploid wheat genotypes is influenced by the environmental factors such as weather, farming practices, soil characteristics or field management, affect how genes are expressed, which may help to explain the situation (Yao *et al*., 2008; Persaud *et al*., 2022). Further investigations will be required to ascertain stability of traits over several years to assess their suitability for crossing with other desirable traits in a breeding programme.

Table 3. 4 Mean squares of combined ANOVA of phenotypic traits of 196 tetraploid wheat accessions evaluated at Sinana and Debrezeit during 2020

Source of Variation	df	SdPSp	SpPSp	SpL	TPP	PH	TKW	GY	Mo	BY	HI	DB	DM
Replication	1	0.44*	115.3***	19.78***	660.5***	9804***	627***	5.21*	5.78***	411***	1898***	64	766***
Genotypes	195	1.05***	6.34***	5.81***	6.18	275.55***	106.94***	5.82***	0.31***	66.67***	684.09***	59.35***	79.81***
Location	1	1.10**	188***	189.29***	220.3***	293536***	10489***	188.54***	56.20***	207***	9488***	5236***	133408***
gen:Loc	195	0.16***	3.01***	1.94*	3	99***	44***	3.39***	0.13***	36.6***	324***	34*	33***
Rep:block	26	0.17*	2.71	2.79**	5.5*	105**	40*	1.59	0.07	17.3	141	54**	47***
Pooled error	365	0.1	1.83	1.22	2.05	49.39	21.25	1.04	0.067	12.39	116.6	14.62	6.12
Mean		3	18	8.2	6	99.86	34.24	3.25	10.62	11.34	31.81	73.4	123.44
SE		0.33	1.42	1.22	1.77	8.13	4.75	1.08	0.29	3.94	10.49	5.12	4.43
CV		12.67	7.65	13.48	23.91	7.04	13.46	31.34	2.66	31.03	33.95	5.21	2
Lsd		0.45	1.88	1.54	1.99	9.77	6.41	1.42	0.39	4.9	15.02	5.32	3.44
Range		1.2-4.6	12.2-23.2	4.0-14	1.0-12.0	53.2-139.2	10.0-52.2	0.06-9.66	10.28-11.73	4.38-22.5	13.55-65.10	53-88	89-154

NB: SdPSp= number of seeds per spikelet, SpPSp = number of spikelet per spike, SpL= spike length (cm), TPP= number of effective tillers per plant, PH= Plant height (cm), TKW= weight of thousand kernels per plant (gm), GY= grain yield (t/ha), BY= biomass yield (t/ha), HI= harvest index, DB= days to booting, DM= days to maturity, SE= standard error, CV = coefficient of variation, Lsd = least signficant difference, gen: Loc = genotype by location interaction, and Rep=block= block within replication.

3.3.2.2. Patterns of quantitative traits variation and performance of the genotypes: The mean performance of the genotypes is given in Annex 3. Four landraces, including 239711, 7060, 5638, and 242789, and two improved varieties, namely Tob66 and Foka, were found to be top performing in their average grain yield; and, their respective average grain yield (t/ha) being 5.98, 4.99, 4.75, 4.72, 4.7, and 4.67. Three of the top performing landraces, 5638, 7060, and 242789, were also among the top performing genotypes in their average biomass yield and were in the acceptable thousand kernel weight range; that is, 7060, 5638, and 239711 recorded high mean value of TKW while 242789 is medium in its TKW. Thirty landraces along with one improved variety, including 5638, 5618, 7003, 7060, 243706,7378,7056,242789, 7242, 7209, 'DZDW176,7,10,DZDW17005,DZDW17006,07', 5586, MCD14-29, 5182, 7626, 5149, MCD3-14, 5441,5913, 5591, 7295, MCD4-32,7063,242781,6914,7664,7825,243717, 'MCD11-7,MCD11-25, MCD13-5, MCD14-19', ETDW/15DZ#044, and Selam, gave a high average value of biomass ranging from 14.4 to 22.5 t/ha (Annex 3) and these genotypes can be potentially produced in areas where wheat straw is used as a source of animal feeds. Sixty three of the genotypes, including sixteen improved varieties and forty seven landraces, had medium thousand kernel weight (36-45); however, 105 of the genotypes were in the low (30-35) kernel weight groups and 28 of the genotypes had very low (<30 gm) kernel weight.

The respective mean and range values for the number of seeds per spikelet (3 and 2-5), the number of spikelets per spike (18 and 15-21), the number of effective tillers per plant (6 and 3-9), spike length (8.20cm and 5.57-10.38cm), plant height (99.86cm and 81.45-117.86cm), thousand kernel weight (34.24g and 23.6-44.9g), grain yield (3.25t/ha and 1.33-5.98t/ha), biomass yield (11.23 t/ha and 3.75-22.5 t/ha), days to booting (73 and 153 61-81), days to maturity (123 and 117-130), and harvest index (33.20% and 13.55-72.21%) are given in Table 3.5. The coefficient of variation was high for grain yield (31.34), biomass yield (31.03), harvest index (33.95), and the number of effective tillers per plant (23.91); intermediate for thousand kernel weight (13.46), spike length (13.48), and the number of seeds per spike (12.67) and low for the number of spikelets per spike (7.65), plant height (7.04), days to booting (5.21), and days to maturity (2) (Table 3.4).

Thousand kernel weight was on average 34.24 gm ranging from 23.6 gm to 44.9 gm, indicating high range values (21.3gm). The minimum thousand kernel weight (23.6 gm) was recorded from the landrace 5166 that was obtained from North Shewa of Amara at an altitude of 2260 m.a.s.l. The maximum thousand kernel weight (44.9gm) was recorded

from the landrace MCD 3-15 that was obtained from East Shewa at an altitude of 2612 m.a.s.l. Grain yield was on average 3.25 (t/ha) with a minimum and maximum average values of 1.33 t/ha and 5.98 t/ha, respectively, and a range of 4.65 t/ha. The minimum grain yield (1.33 t/ha) was recorded from the landrace ETDW/15DZ014 obtained from West Shewa at an altitude of 2597 m.a.s.l. The maximum grain yield (5.98 t/ha) was recorded from the landrace 239711 obtained from Bale at an altitude of 2760 m.a.s.l. Moisture content was on average 10.62 % with a minimum and maximum average value of 10.16 and 11.73, respectively, and a range of 10.62%.

Table 3. 5 The mean, minimum, maximum, and range values of quantitative traits at the entire genotypic level.

	Traits										
Values	PH	TPP	SpSp	SDpSp	SPL	TKW	GY	HI	BY	DB	DM
Mean	99.86	6	18	3	8.2	34.24	3.25	33.2	11.23	73	123
Min	81.45	3.32	15	2	5.57	23.6	1.33	13.55	3.75	61	117
Max	117.86	8.5	21	4	10.38	44.9	5.98	63.72	22.5	81	130
Range	36.41	5.18	6	2	4.81	21.3	4.65	50.17	18.75	20	13

Remarks: SDpsp= the number of seeds per spikelet, SpPSp = the number of spikelet per spike, SPL= spike length (cm), TPP= the number of effective tillers per plant, PH= Plant height (cm), TKW= weight of thousand kernels per plant (gm), GY= grain yield (t/ha), BY= biomass yield (t/ha), HI= harvest index, DB= days to booting, DM= days to maturity, Min= minimum, and Max= maximum.

The minimum seed per spikelet was for 239711 from Bale at an altitude of 2760 m.a.s.l. while maximum seed per spikelet was recorded from landrace 7673 from east Gojam at altitude of 2560 m.a.s.l. Spikelet per spike, minimum value (15) was recorded for landrace 7647 from east Gojam at an altitude of 2470 m.a.s.l. while the maximum number of spikelet per spike (21) was obtained from landrace 5638 at an altitude of 2160 m.a.s.l. from North Shewa. Minimum number of effective tillers per plant (3) was recorded from landrace 7191 at an altitude of 2500 from North Shewa while the maximum number of effective tillers per plant (9) was obtained from landrace 5149 at an altitude of 2330 from West Shewa. Spike length; minimum value recorded was 5.57 cm from landrace 7690 at an altitude 2840 m.a.s.l. of from West Gojam while the maximum spike length was 10.34cm from landrace 5581 obtained from North Shewa at altitude of 2520 m.a.s.l.

Minimum plant height (81.45 cm) was recorded for the landrace MCD4-13, MCD8-6, MCD11-18 at an altitude of 2671 m.a.s.l. while the maximum height was 117.86 cm from landrace MCD10-3, MCD10-16, MCD10-44 at an altitude of 2484 m.a.s.l.

Biomass yield minimum biomass obtained was 3.75(t/ha) from landrace 7015 at an altitude of 2395 from Arsi while the maximum biomass yield was 22.5(t/ha) from landrace 7295

South Gondor at an altitude of 2640. Earliest days to booting was 61 for the landrace 7664 at an altitude of 2430 m.a.s.l. from East Gojam while the late days to booting was 81 from an improved cultivar Quamy. Earliest date was 117 from landrace 7664 obtained from East Gojam at an altitude of 2430 m.a.s.l. while late maturity (130) was from landrace ETDW/15DZ#43ETDW/15DZ#068 at an altitude of 2461 m.a.s.l.

3.3.2.3. Comparison of selected genotypes with the original population: The performance of the top 5% of genotypes is compared with the population mean and the mean of the released varieties (Table 3.6). Many other superior genotypes for multiple traits of agronomic significance, including grain yield, were identified. The best genotypes for grain yield (t/ha) were found to be superior by 40.31% - 83.69% to the population mean and 28.81% -68.64% to the mean of grain yield performance of the released varieties; plant height (cm) 10.10% - 18.03% to the population mean and 18.12% - 26.62% to the mean plant height performance of the released varieties; thousand kernel weight (g) 20.04%-31.13% to the population mean and 9.28%-19.38% to the mean thousand kernel weight performance of the released varieties; biomass yield (t/ha) 39.14% - 100.4% to the population mean and 50.68% - 117% to the mean biomass performance of the released varieties; days to booting 7.97% - 10.69% to the population mean and 15.27% - 18.18% to the mean days to booting of the released varieties; days to maturity 3.7% - 5.32% to the population mean and 5.04%-6.68% to the mean days to maturity of the released varieties; spikelets per spike 8.25%-17.24% to the population mean and 12.71% - 22.07% to the mean performance of the number of spikelets per spike of the released varieties, and the number of seeds per spikelet 50.60%-80.72% to population mean and 2.74%-23,29% to the mean performance of the number of seeds per spikelet of the released varieties (Table 3.6).

It was evident that most genotypes of tetraploid wheat showed superior phenotypic performances when compared to the population mean and the mean released varieties (Figure 3.4).

The majority of the genotypes outperformed the mean of released varieties with respect to the number of spikelets per spike, spike length (cm), plant height (cm), the number of effective tillers per plant, days to booting, and days to maturity; however, most of the genotypes outperformed the mean of the population with respect to spike length (cm), plant height (cm), and days to booting.

The numbers of genotypes outperforming the population mean were higher than the numbers of genotypes outperforming the mean of released varieties with respect to the number of seeds per spikelets, thousand kernel weight (g), grain yield (t/ha), and harvest index. On the other hand, the numbers of genotypes outperforming the mean of released varieties were higher than the numbers of genotypes outperforming the population mean with respect to the number of spikelets per spike, spike length (cm), plant height (cm), the number of effective tillers per plant, days to booting, and days to maturity (Figure 3.4). Overall, there are several genotypes showing superior performance over the population mean and the mean of released varieties.

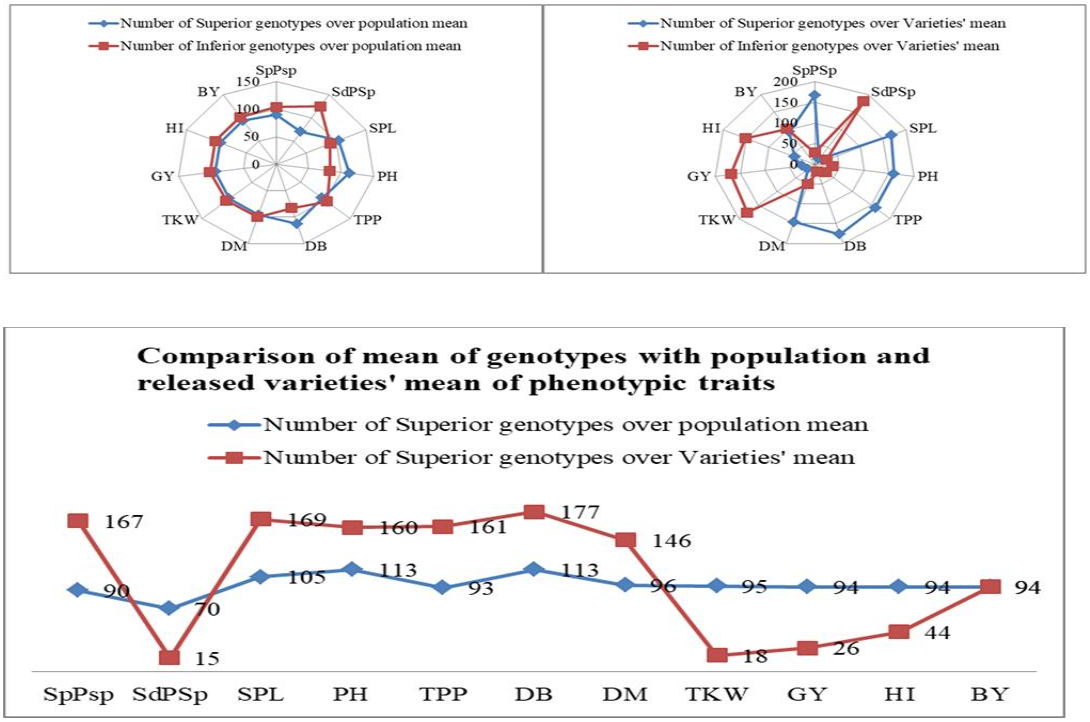


Figure 3. 4 The number of superior and inferior tetraploid wheat genotypes over (A) the population means, (B) the mean of released varieties, and (C) comparison of number of superior genotypes over population mean and released varieties' mean of phenotypic traits. Remarks: BY= biomass yield (t/ha), HI = harvest index, GY = grain yield (t/ha), TKW= thousand kernel weight (g), DM= days to maturity, DB= days to booting, TPP= number of effective tillers per plant, PH= plant height (cm), SPL= spike length (cm), SdPSp= number of seeds per spikelet, and SpPSp= number spikelets per spike.

Table 3. 6. Comparison of mean performance of top 5% tetraploid wheat genotypes with mean of the population and the released varieties for phenotypic traits of tetraploid wheat grown at Sinana and Debrezeit during 2020

Genotypes	Mean	Relative advantage (%) over		Genotypes	Mean	Relative advantage (%) over		Genotypes	Mean	Relative advantage (%) over		Genotypes	Mean	Relative advantage (%) over	
		MP	MRV			MP	MRV			MP	MRV			MP	MRV
	Plant height			Thousand kernel weight(g)				Grain yield (t/ha)				Biomass yield (t/ha)			
5052	117.86	18.03%	26.62%	MCD3-15	44.9	31.13%	19.38%	239711	5.97	83.69%	68.64%	5638	22.5	100.4%	117%
6975	113.05	13.21%	21.45%	1	44.65	30.4%	18.72%	7060	4.99	53.54%	40.96%	5618	19.38	72.53%	86.84%
242793	112.65	12.81%	21.02%	242781	43.7	27.63%	16.19%	5638	4.75	46.15%	34.18%	7003	18.13	61.40%	74.78%
7060	111.75	11.91%	20.06%	Selam	43.3	26.46%	15.13%	242789	4.72	45.23%	33.33%	7060	17.5	55.83%	68.76%
243706	111.65	11.81%	19.95%	Gerardo	42.35	23.69%	12.60%	Tob66	4.7	44.62%	32.77%	243706	16.88	50.27%	62.73%
5760	111.6	11.76%	19.90%	Quamy	41.9	22.37%	11.41%	Foka	4.67	43.69%	31.92%	7378	16.25	44.70%	56.70%
7673	111.55	11.71%	19.84%	Foka	41.75	21.93%	11.01%	7015	4.62	42.15%	30.51%	7056	16.25	44.70%	56.70%
Boahi	111.5	11.66%	19.79%	6914	41.55	21.35%	10.48%	6914	4.6	41.54%	29.94%	242789	15.63	39.14%	50.68%%
5140	110.3	10.45%	18.50%	Tob66	41.5	21.20%	10.34%	7641	4.58	40.92%	29.38%	7242	15.63	39.14%	50.68%
242781	109.95	10.10%	18.12%	7762	41.1	20.04%	9.28%	7056	4.56	40.31%	28.81%	7209	15.63	39.14%	50.68%
MP	99.86		7.28	MP	34.24	0	-8.96	MP	3.25	0	-8.19	MP	11.23	0	8.29
MRV	93.08	-6.79		MRV	37.61	9.84	0	MRV	3.54	8.92	0	MRV	10.37	-7.66	0
	Days to booting			Days to maturity				Spikelets per spike				Number of seeds per spikelet			
5369	81.25	10.69%	18.18%	5584	130	5.32%	6.68%	5158	20.74	17.24%	22.07%	Alemtena	4.5	80.72%	23.29%
7762	80.5	9.67%	17.09%	7464	129.5	4.92%	6.27%	7046	20.45	15.60%	20.36%	Fetan	4.3	72.69%	17.81%
2	80.25	9.33%	16.73%	5586	129.5	4.92%	6.27%	3	19.9	12.49%	17.13%	DZDW17002	4.1	64.66%	12.33%
243703	80	8.99%	16.36%	MCD8-44	129	4.51%	5.86%	242781	19.58	10.68%	15.24%	DZ2018	4	60.64%	9.59%
7581	79.75	8.65%	16%	7477	128.8	4.31%	5.65%	7	19.5	10.23%	14.77%	Malefia	4	60.64%	9.59%
7664	79.75	8.65%	16%	5	128.5	4.11%	5.45%	5600	19.5	10.23%	14.77%	Tob66	3.9	56.63%	6.85%
5504	79.75	8.65%	16%	Tob66	128.5	4.11%	5.45%	5434	19.3	9.10%	13.60%	Selam	3.8	52.61%	4.11%
4	79.5	8.31%	15.64%	5465	128.5	4.11%	5.45%	8	19.25	8.82%	13.30%	Werer	3.8	52.61%	4.11%
5492	79.5	8.31%	15.64%	6	128.3	3.91%	5.24%	5526	19.2	8.54%	13.01%	Cocorit/71	3.8	52.61%	4.11%
5526	79.25	7.97%	15.27%	243717	128	3.70%	5.04%	7063	19.15	8.25%	12.71%	Ejersa	3.75	50.60%	2.74%
MP	73.4	0	6.76	MP	123.4	0	1.29	MP	17.69	0	4.12	MP	2.49	0	-31.78
MRV	68.75	-6.34	0	MRV	121.9	-1.27	0	MRV	16.99	-3.96	0	MRV	3.65	46.59	0

1= ETDW/15DZ#044, 2 = ETDW/15DZ036, ETDW/15DZ063, 3 = MCD13-8, MCD14-29, 4 = DZDW17004,DZDW1702,04, 5 = DZDW17004,DZDW1702,04, 6 = ETDW/15DZ#4,ETDW/15DZ#20……..,7 = MCD5-4, MCD5-25, MCD6-44,……, and 8 = ETDW/15DZ036,ETDW/15DZ063

A student t-test was used to compare the mean of the top 5% genotypes to the mean of the population for all phenotypic traits (Table 3.7). The t-test showed highly significant ($p \leq 0.0001$) differences between means of the selected the top 5% best genotypes ($\bar{x}$) and the population parameters (μ) for plant height (cm), the number of effective tillers per plant, the number of spikelets per spike, the number of seeds per spikelet, spike length (cm), thousand kernel weight (g), grain yield (t/ha), harvest index, biomass yield (t/ha), days to booting, and days to maturity. The study revealed that the top 5% of genotypes had greater relative advantages in all phenotypic traits, with differences between them and the population mean performances ranging from 4.39 % for days to maturity to 91.75 % harvest index. The top 5% genotypes gave 48.31% greater grain yield (t/ha) and 54.76 % biomass yield (t/ha) than the population as whole. Additionally, the top 5% of genotypes exhibited advantages of 12.35, 34.56, 11.14, 60.44, 21.61, 24.62, 54.76, and 8.72 % in plant height (cm), the number of effective tillers per plant, the number of spikelets per spike, the number of seeds per spikelet, spike length (cm), thousand kernel weight (g), harvest index, biomass yield (t/ha), and days to booting above the mean performances of the population, respectively, indicating the occurrence of various degrees of tetraploid wheat enhancements through selection.

Table 3. 7 Comparison of mean performances of selected top 5% genotypes with the population mean performances for phenotypic traits of tetraploid wheat germplasm grown at Sinana and Debrezeit during 2020

Traits	Mean of selected genotypes	Population parameter (μ)	Change through selection ($\bar{X}$-μ)	Change as % of a population parameter (μ)	t-test
PH	112.19	99.86	12.33	12.35	17.75***
TPP	8.06	5.99	2.07	34.56	25.06***
SpPSp	19.66	17.69	1.97	11.14	11.44***
SdPSp	3.995	2.49	1.505	60.44	19.45***
SpL	9.972	8.2	1.772	21.61	29.02***
TKW	42.67	34.24	8.43	24.62	19.41***
GY	4.82	3.25	1.57	48.31	11.68***
HI	63.66	33.2	30.46	91.75	19.31***
BY	17.38	11.23	6.15	54.76	8.92***
DB	79.95	73.4	6.55	8.92	35.29***
DM	128.85	123.43	5.42	4.39	26.91***

The tabulated t-value for 9 degrees of freedom is 4.781 for two-tailed tests and 4.297 for one-tailed tests at 0.1% probability. T = Student t-test; *** = significantly different at 0.1% probability level. NB: PH=plant height, DB= days to booting, DM= days to maturity, TKW= thousand kernel weight, TPP= number of effective tillers per plant, GY= grain yield (t/ha), HI= harvest index, BY= biomass yield (t/ha), SpL= spike length (cm), SdPSp= seed per spikelet, and SpPSp= spikelet per spike.

3.3.2.4. Phenotypic and genotypic coefficient of variation

Phenotypic coefficient of variation is a measure of variation due to genetic and environmental factors and genotypic coefficient of variation (GCV) is a measure of the relative variability of a trait due to genetic differences among individuals. The values of phenotypic coefficient of variation (PCV = the variation due to genotype and environment), genotypic coefficient of variation (GCV = the variation due genotype only), genotype by environment interaction coefficient of variation (GECV = the variation due to the interaction of genotype and environment), broad sense heritability (H^2 = how much a variation in a trait is due to genetic factors), and genetic advance (explains the degree of gain obtained in a trait under a particular selection pressure) are given in Table 3.8. PCV and GCV below 10%, 10-20%, and above 20% were respectively regarded as low, intermediate, and high (Burton and DeVane 1953). The values of PCV and GCV were low for the number of spikelets per spike (6.98, 5.06), plant height (8.31, 6.65), moisture content (2.66, 2.11), days to booting (5.25, 3.43), and days to maturity (3.62, 2.77) indicating minimum variability in these traits. The low PCV, GCV, and ECV for the number of spikletes per spike, plant height, days to booting, and days to maturity implies that the improvement of these traits by selection is not possible as there is inadequate heritable variation among the genotypes in spite of the low influence of environment (Joshi *et al*., 2018). The values of PCV and GCV were intermediate for the number of seeds per spikelet (17, 15.64), spike length (14.74, 12.01), and thousand kernel weight (15.1, 11.58). The values of PCV and GCV were high for grain yield (37.18, 24.03), biomass yield (36, 24.18), and harvest index (41.11, 29.82) indicating presence of high variability in these traits. High PCV (20.79) and intermediate GCV (14.93) values were obtained for the number of effective tillers per plant.

Differences between the value of GCV and PCV respectively considered high, moderate, and low if more than 20%, 10–20%, and less than 10% (Deshmukh *et al*., 1986). The difference between PCV and GCV was 0.85, 1.36, 1.66, 1.82, 1.92, 2.73, 3.52, 5.86, 11.28, 11.82.13.15 for days to maturity, number of seeds per spikelet, plant height, days to booting, number of spikelets per spike, spike length, thousand kernel weight, number of effective tillers per plant, harvest index, biomass yield, and grain yield respectively. The difference between PCV and GCV was moderate for grain yield, biomass yield, and harvest index and low for days to maturity, number of spikelets per spike, plant height, days to booting, number of seeds per spikelet, spike length, thousand kernel weight, and

number of effective tillers per plant implying that grain yield, biomass yield, and harvest index were influenced by the environment whereas the remaining traits were mainly due to genetic factors and there is negligible influence on the expression of these traits by the environment (Prasad *et al*., 2023). Similar to this study, Arega *et al.* (2010) reported low values of PCV and GCV in days to maturity (4.1, 3.6) and plant height (9.7, 9.1) and intermediate PCV and GCV in spike length (13.2, 12.5); however, their result disagrees with the present result on the number of effective tillers per plant, the number of spikelets per spike, biomass yield, thousand kernel weight, harvest index, and grain yield. Additionally, similar to this study, Abebe and Desta (2017) reported low PCV and GCV in days to maturity (5.73, 5.15), high PCV and intermediate GCV in the number of effective tillers per plant (20.82, 16.82), and high PCV and GCV in biomass yield (24.98, 23.67) and harvest index (25.97, 25.72); however, their work disagrees with the present study on spike length, grain yield, and plant height. Moreover, similar to this study, Azene *et al.* (2020) reported low PCV and GCV in days to maturity (4.94, 3.08), intermediate PCV and low GCV in spikelet per spike (11.75, 9.63) and spike length (11.91, 6.82), and intermediate PCV and GCV in thousand kernel weight (14.64, 12.67), and; however, their result disagree with the present study on the GCV and PCV values of other traits. Furthermore, in line with this study, Meles *et al.*(2017) reported low PCV and GCV values for plant height (8.08, 6.45) and days to maturity (7.99, 7.48) and intermediate PCV and low GCV in thousand kernel weight (11.04, 8.45); however, their result disagrees with the present result on effective tillers per plant, the number of spikelet per spike, spike length, the number of effective tillers per plant, thousand kernel weight, harvest index, grain yield, and biomass yield. Similar to this study Belay *et al*. (1993) reported intermediate PCV and GCV in TKW (10.90, 10.22), low PCV and GCV in the number of spikelets per spike (7.84, 7.08), in the number of seeds per spikelet (9.7, 9.11); in plant height (7.66, 6.52), and in days to maturity (3.09, 3.00) and high PCV in grain yield (21) and in the number tillers per plant (24); however, opposite to the present study they reported intermediate PCV and GCV in biomass yield (14.44, 11.60) and in harvest index (14.32, 12.42), intermediate GCV in grain yield (18.60), and high GCV in the number tillers per plant (22.36).

The observed environmental coefficient of variation (ECV) was high for grain yield (31.38), biomass yield (31.04), harvest index (33.95), and the number of effective tillers per plant (23.9) implying that the variation among genotypes is not heritable and so, selection is not effective in the improvement of these traits (Joshi *et al*., 2018);

intermediate for the number of seeds per spikelet (10.54), spike length (13.47), and thousand kernel weight (13.46); and low for the number of spikelets per spike (7.52), plant height (7.04), moisture content (2.49), days to booting (4.24), and days to maturity (2). In line with this study, Joshi *et al*. (2018) reported high values of environmental coefficient of variation for the number of effective tillers per plant (26.146) and low number spikelets per spike (8.83), plant height (4.755), and days to maturity (1.209).

The value of genotype by environment interaction variation was high for grain yield (33.42), biomass yield (30.69), and harvest index (32.01) implying the performance of these traits varies across locations and that there is effect of genotypes, environment had interaction of genotype and environment in the expression of these traits; intermediate for the number of effective tillers per plant (11.57) and low for the number of seeds per spikelet (5.77), the number of spikelets per spike (4.27), spike length (7.32), plant height (4.99), thousand kernel weight (9.85), moisture content (1.63), days to booting (4.24), and days to maturity (2.97) showing these traits are stable across locations and that the influence of the environment is minimal on these traits.

Table 3. 8 Variability, Heritability and Genetic Advance of tetraploid wheat germplasm of Ethiopia based on quantitative traits

	Traits											
Parameters	SdPSp	SpPSp	SpL	TPP	PH	TKW	GY	Mo	BY	HI	DB	DM
GA	0.22	0.83	0.97	0.8	44.14	15.73	0.61	0.05	7.52	90.02	6.34	11.7
EV	0.1	1.83	1.22	2.05	49.39	21.25	1.04	0.07	12.39	116.6	14.62	6.12
GIEV	0.03	0.59	0.36	0.48	24.81	11.38	1.18	0.03	12.11	103.7	9.69	13.44
PV	0.26	1.58	1.46	1.55	68.89	26.73	1.46	0.08	16.67	171.02	14.84	19.95
GCV	15.64	5.06	12.01	14.93	6.65	11.58	24.03	2.11	24.18	29.83	3.43	2.77
PCV	17	6.98	14.74	20.79	8.31	15.1	37.18	2.66	36	41.11	5.25	3.62
ECV	10.54	7.52	13.47	23.9	7.04	13.46	31.38	2.49	31.04	33.95	5.21	2
GECV	5.77	4.27	7.32	11.57	4.99	9.85	33.42	1.63	30.69	32.01	4.24	2.97
H^2	84.62	52.53	66.44	51.61	64.07	58.85	41.78	62.5	45.11	52.64	42.72	58.65
GA	88.71	135.76	165.06	132.11	1093.34	625.56	103.79	36.35	378.67	1415.35	338.35	538.6
GAM	29.57	7.54	20.13	22.06	10.95	18.27	31.94	3.42	33.39	44.49	4.61	4.36
PCV-GCV	1.36	1.92	2.73	5.86	1.66	3.52	13.55	0.55	11.82	11.28	1.82	0.85

NB: PH=plant height, DB= days to booting, DM= days to maturity, TKW= thousand kernel weight, TPP= number of effective tillers per plant, GY= grain yield (t/ha), HI= harvest index, BY= biomass yield (t/ha), SpL= spike length (cm), SdPSp= seed per spikelet, SpPSp= spikelet per spike, GA=Genetic Variance, EV=Environmental Variance, GIEV=Genotype by Environment interaction variance, PV=Phenotypic Variance, GCV=Genotypic Coefficient of Variation, PCV=Phenotypic Coefficient of Variation. ECV=Environmental Coefficient of Variation, GECV=Genotype by Environment interaction Coefficient of Variation, H^2=Broad sense heritability, GA=Genetic Advance, and GAM=Genetic Advance as a Percentage of Mean.

3.3.2.5. Broad sense heritability (H^2) and genetic advance

H^2 values <40%, 40-80%, and > 80% were categorized as low, medium, and high, respectively (Mesele *et al*., 2015). Estimates of heritability (H^2) ranged from 41.78% to 84.62% for grain yield and the number of seeds per spikelet, respectively (Table 3.8). High value of broad sense heritability was observed for the number of seeds per spikelet (84.62) and medium value of heritability was recorded for spike length (66.44), plant height (64.07), days to booting (42.72), days to maturity (58.65), the number of spikelets per spike (52.53), the number of effective tillers per plant (51.61), thousand kernel weight (58.85), biomass yield (45.11), grain yield (41.78), and harvest index (52.64). The high heritability estimates for the number of seeds per spikelets indicated that strong genetic nature for the trait implying that selection for this trait might be effective in this set of genotypes (Prasad *et al*., 2023). Similar to this study, Belay *et al.* (1993) reported high H^2 for the number seeds per spikelets (88.2%) and medium H^2 for biomass yield (64.6%); however, their report for other traits disagrees with the present study. In line with this study, Meles *et al.*(2017) reported medium heritability in plant height (63.99), number of spikelets per spike (55.76), spike length (46.73), biomass yield (45.29), grain yield (52.83), and TKW (57.41); however, their report on other traits disagrees with this study. In line with this study, Arega *et al*. (2010) reported medium H^2 in days to maturity (71.7), number of effective tillers per plant (42.9), number of spikelets per spike (75.2), biomass yield (53.5), grain yield (47.1), and harvest index (71.2); however, their report in other traits on H^2 disagrees with the present study.

Heritability and genetic advance decide the success of selection in breeding programs. The values of genetic advance as a percentage of mean >20%, 10-20%, and <10% were categorized as high, intermediate, and low, respectively (Johnson, *et al*., 1955). Genetic advance as the percentage of mean was low for the number of spikelet per spike (7.54%), days to booting (4.61%), and days to maturity (4.36%); intermediate for plant height (10.95) and thousand kernel weight (18.27), and high for spike length (20.13), the number of seeds per spikelet (29.57), the number of effective tillers per plant (22.06), grain yield (31.94), biomass yield (33.39), and harvest index (44.49). The observed genetic advance and broad sense heritability (H^2) were high for the number of seeds per spikelet implying that the trait is governed by additive gene action and that selection may be effective for this specific trait (Prasad *et al*., 2023). In line with this study, Belay *et al*. (1993) reported low GAM (6) in days to maturity, intermediate GAM in plant height (11.4),TKW (19.8), and

high GAM in number of tillers per plant (42.9) and harvest index (25) and intermediate heritability in plant height (72.3) and biomass yield (64.9); however, in contrast they reported high heritability in days to maturity (94.5), number of tillers per plant (86.7), number of spikelets per spike (81.4), number of seeds per spikelets (88.2), in thousand kernel weight (88), and harvest index (87) and intermediate values of GAM in TKW, biomass yield, and number of seeds per spikelet (17.6). The work by Arega *et al*. (2010) on durum wheat Ethiopia agrees with the present study on GAM values of plant height (17.4), thousand kernel weight, and days to maturity while it disagrees with this study for the other traits. Moreover, Mesele *et al.*(2015) reported similar result on biomass yield, grain yield, spike length, thousand kernel weight, and plant height in bread wheat of Ethiopia; however, their report disagrees with the present study on the number of effective tillers per plant, harvest index, seed per spikelet, and days to maturity. Furthermore, Azene *et al.* (2020) also reported similar to the present study on days to maturity, harvest index, and thousand kernel weight; however, their report disagrees with present study on plant height, spikelet per spike, spike length, biomass yield, and grain yield. Similar to this study, Meles *et al.*(2017) reported intermediate GAM for plant height (10.64), TKW (13.08) and low GAM in the number of spikelets per spike (9.99); however, their report on other traits disagrees with this study.

The number of seeds per spikelet, the number of effective tillers per plant, grain yield, biomass yield, and harvest index gave high values of genetic advance as a percentage of mean. Heritability estimates with genetic advance are more useful in predicting the gain under selection than heritability alone. High heritability accompanied with high genetic advance is an indication of additive gene effects (Johnson *et al*., 1955) and hence, high genetic gain from selection of the number of seeds per spikelet would be expected. High heritability along with high expected genetic gain could result from the high additive gene effect and therefore targeted selection for the number of seeds per spikelet will lead to yield improvement (Panse and Sukhatme 1957).

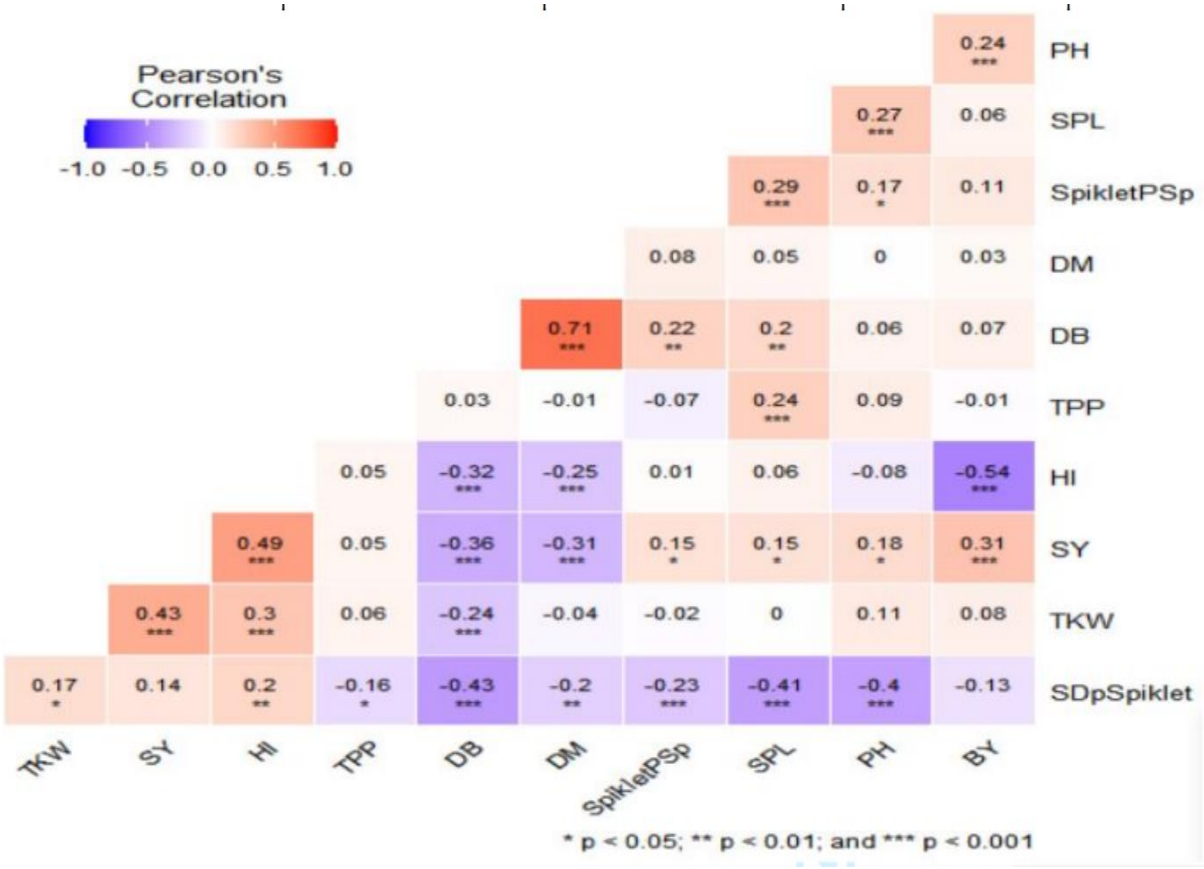
In general, the tetraploid wheat genotypes studied showed highly heritable variation with high expected genetic advance for the number of seeds per spikelet. Thus, the top 5% of the tetraploid wheat genotypes for the number seeds per spikelet were DZDW17002, Alemtena, Fetan, Malefia, DZ2018, Tob66, Werer, Selam, Cocorit/71, and Ejersa. Therefore, the tetraploid wheat genotypes selected for the greater number of seeds per

spikelet could be used as potential parents for crossing to develop improved varieties for this trait.

3.3.2.6. Correlation analysis

The result of Pearson correlation coefficient is given in Table 3.9. Grain yield showed a highly significant ($p<0.001$) negative correlation with days to booting (-0.36***) and days to maturity (-0.31***) and a highly significant ($p<0.001$) positive correlation with thousand kernel weight (0.43***), biomass yield (0.31***), and harvest index (0.49***). Grain yield, on the other hand, showed a significant (0.01) positive correlation with plant height (0.18*), spike length (0.15*), and the number of spikelets per spike (0.15*), and a positive correlation with the number of seeds per spikelet (0.14), and the number of effective tillers per plant (0.05). Grain yield showed positive association with thousand kernel weight, biomass yield, harvest index, plant height, spike length, the number of spikelets per spike, seeds per spikelet, and effective tillers per plant. This implies that there might be common gene(s) that control grain yield and these traits, which indicates that improving either one or more of these traits could result in high grain yield (Arega *et al.*, 2010). According to Kearsey and Pooni (1996), the positive association of these traits with grain yield might be due to either the presence of strong coupling of genes or pleiotropic genes controlling the traits in the same direction. In line with this study, Belay *et al*. (1993) reported highly negative association between days to maturity and grain yield and highly significant positive association of grain yield with biomass yield, harvest index, thousand kernel, number of tillers per plant, number of seeds per spikelet, and number of spikelets per spike.

Table 3. 9 Pearson's correlations of quantitative traits of 174 tetraploid wheat landraces and 22 improved cultivars of Ethiopia



Remarks: TKW= thousand kernel weight, SY= Seed yield (t/ha), HI= harvest index, TPP= number of effective tillers per plant, DB= days to booting, DM= days to maturity, SpikletPSp = number of spikelet per spike, SPL= spike length, PH= plant height, and BY= biomass yield.

In line with the present study, Arega *et al.*(2010) reported that grain yield had significant association with biomass yield, plant height, thousand kernel weight, and harvest index at both phenotypic and genotypic level. This result disagrees with the work of Baye *et al.* (2020). Similarly, Azene *et al.* (2020) reported that there was a highly significant positive correlation of grain yield with thousand kernel weight, biomass yield, and harvest index and a positive correlation with spike length. Grain yield showed a highly significant ($p<0.001$) negative correlation with days to booting (-0.36***) and days to maturity (-0.31***). In line with this study, Ayer *et al.* (2017) reported non-significant negative correlation of days to booting and days to maturity with grain, highly significant positive correlation of grain yield with spike length, highly significant negative correlation with 1000 grain weight, and harvest index. Similar to this study, Sharma *et al.* (2017) also

reported highly significant (-0.489**) negative correlation of days to maturity with grain yield in wheat.

3.3.2.7. Principal component analysis

At the whole genotype level, the first and the second principal components, respectively, accounted for 19.74 % and 15.96% of the total variation (Table 3.10; Figure 3.5). The third, the fourth, the fifth, and the six principal components, respectively, accounted for 9.53, 8.07, 7.91, and 6.38 of the total variation. The first five principal components, with eigenvalues greater than one, together explained 61.21% of the total variation. The variation in the first principal component was mainly contributed by days to booting, the number of seeds per spikelet, days to maturity, spike density, and harvest index. Grain yield, spike length, and plant height were the traits that contributed to the variation in the second principal component. Biomass yield and harvest index contributed to the variation in the third principal component. The work of Azene *et al.*(2020) similarly indicated that days to maturity contributed to the variation in the first component and grain yield to the second principal component. Similar to this study, Hailu *et al.*(2006) reported that spike density contributed to the first component in tetraploid wheat of Ethiopia.

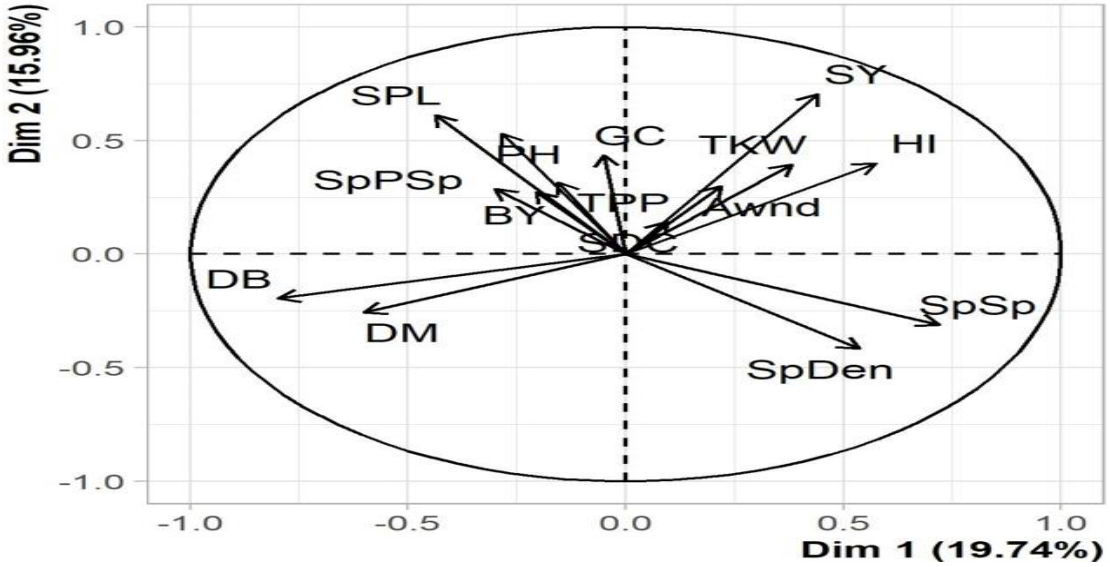


Figure 3. 5 PCA Biplot of whole genotypes of tetraploid wheat of Ethiopia based on 15 phenotypic traits

Remarks: SDpSp= the number of seeds per spikelet, SpPSp = the number of spikelets per spike, SPL= spike length (cm), TPP= the number of effective tillers per plant, PH= Plant height (cm), TKW= weight of thousand kernels per plant (gm). SY= seed yield (t/ha), BY= biomass yield (t/ha), HI= harvest index, DB= days to booting, DM= days to maturity, GC= glume color, SpDen=spike density, SDC= seed color, and Awnd=awndness.

Table 3. 10 Principal components analysis of 15 phenotypic traits of tetraploid wheat genotypes

Traits	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15
PH	2.72444	11.7326	4.08503	0.16805	4.16593	1.60401	5.95371	0.00047	48.6804	4.63994	0.39455	0.43902	15.0748	0.3253	0.01184
TPP	0.79885	4.16896	6.87719	6.78258	10.4857	0.00333	19.5182	19.5179	8.69613	21.3205	1.1123	0.60914	0.07897	0.02055	0.00967
SpPSp	3.04704	3.42528	1.23172	23.1407	10.4451	13.9360	4.69392	0.30395	5.41854	16.7258	1.78831	15.6267	0.00244	0.19973	0.01477
SDpSp	17.4749	4.03039	0.42805	0.66976	5.33674	0.00145	3.67391	1.21251	1.90311	0.28116	1.36334	0.68114	58.1257	4.64883	0.16897
SPL	6.4591	15.5981	3.76971	0.00390	0.00925	3.76706	2.68909	1.28579	0.61022	2.00493	22.4776	35.9415	4.53147	0.62760	0.22473
TKW	4.91129	6.49101	0.90829	15.0948	5.22869	12.0445	12.2737	0.07044	0.00606	2.39973	29.0955	3.28328	4.87665	3.15175	0.16442
BY	1.43311	3.11283	51.5858	1.15624	3.99089	0.28877	0.06319	0.10514	5.24159	1.59180	1.77652	0.38637	0.07746	0.01959	9.17073
HI	11.1114	6.66734	20.2989	9.64187	0.49800	0.16429	0.09420	0.16118	0.74624	0.09888	11.3249	0.60334	0.21580	0.00136	8.37239
GY	6.55144	20.7495	3.45255	4.64364	0.07630	0.02128	0.08173	0.54914	2.74397	2.70126	23.0886	3.47708	0.14433	0.42337	1.29580
DB	21.5553	1.58593	0.25698	8.66216	1.10803	3.33670	0.37802	0.01973	0.48928	0.43099	5.37625	4.46582	0.01397	52.2275	0.09335
DM	12.1882	2.73112	0.08261	14.3352	8.00774	20.3021	1.22715	0.09472	2.2E-06	0.13309	0.91099	0.00934	3.16999	36.7956	0.01220
GC	0.08350	7.87125	0.68647	6.32255	0.03425	6.34746	0.45315	59.5936	9.50626	7.04440	0.02384	0.25692	1.72178	0.00467	0.04994
SpDen	9.74446	7.27806	5.94518	2.81467	2.16686	0.40682	2.40115	0.40000	0.02487	31.4936	0.55138	33.8863	2.26790	0.61773	0.00100
SDC	0.30299	0.82893	0.00011	0.12223	47.2052	12.7755	9.60074	2.81088	15.7229	5.30358	0.00862	0.00905	4.89180	0.02148	0.39608
Awnd	1.61398	3.72871	0.39148	6.44170	1.24132	25.0008	36.8982	13.8746	0.21049	3.83043	0.70734	0.32500	4.80697	0.91501	0.01411
EV	2.96086	2.39362	1.42944	1.21062	1.18673	0.95657	0.94624	0.89366	0.74627	0.65117	0.52903	0.42483	0.39000	0.21534	0.06563
PV	19.74	15.96	9.53	8.07	7.91	6.38	6.31	5.96	4.98	4.34	3.53	2.83	2.60	1.44	0.44
CV	19.74	35.70	45.23	53.30	61.21	67.59	73.89	79.85	84.83	89.17	92.69	95.53	98.13	99.56	100

Remarks: SDpSp= the number of seeds per spikelet, SpPSp = the number of spikelets per spike, SPL= spike length (cm), TPP= the number of effective tillers per plant, PH= Plant height (cm), TKW= weight of thousand kernels per plant (gm). GY= grain yield (t/ha), BY= biomass yield (t/ha), HI= harvest index, DB= days to booting, DM= days to maturity, GC= glume color, SpDen=spike density, SDC= seed color, Awnd=awndness, EV= eigenvalue, PV= percentage of variance, and CV = cumulative variance.

Regions of collections based PCA

Principal component analysis based on regions of collections, altitudinal classes, and species level is given Table 3.11. Based on the regions of collections, variances of 69.9 %, 20.8 %, and 8.0 % were extracted from the first three components with eigenvalues greater than one and the first two components explained 90.7 % of the total variation (Table 3.11; Figure 3.6). Numbers of seeds per spikelet, the number of spikelets per spike, thousand kernel weight, days to booting, spike length, spike density, awndness, biomass yield, and glum color were the major contributors to the variation in the first principal component. Grain yield, harvest index, days to maturity, and number of effective tillers per plant majorly contributed to the variation in the second principal component.

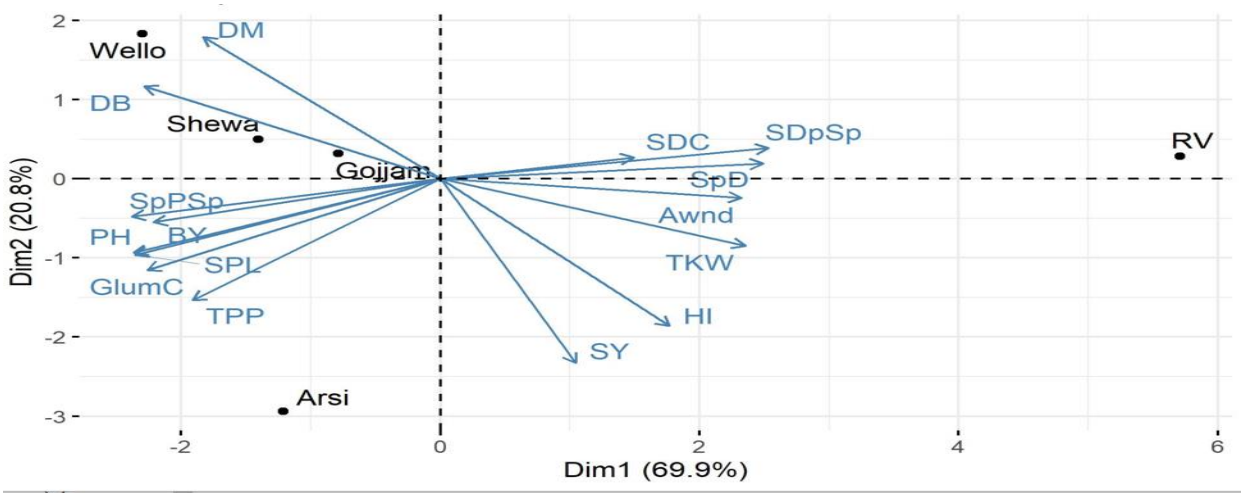


Figure 3. 6 PCA- Biplot of regions of collection of tetraploid wheat genotypes based on the fifteen phenotypic traits. Remarks: SDpsp= the number of seeds per spikelet, SpPSp = the number of spikelets per spike, SPL= spike length (cm), TPP= the number of effective tillers per plant, PH= Plant height (cm), TKW= weight of thousand kernels per plant (gm). GY= grain yield (t/ha), BY= biomass yield (t/ha), HI= harvest index, DB= days to booting, DM= days to maturity, GC= glume color, SpDen=spike density, SDC= seed color, and Awnd=awndness.

Table 3. 11 Principal component analysis of the 15 phenotypic traits of tetraploid wheat based on the species level, the regions of collections, and the altitudinal classes

Traits	Regions of collection				Altitudinal classes							Species		
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC1	PC2	PC3
PH	8.09	4.276	1.3967	0.9392	5.745	8.754	5.2485	6.047	24.216	28.7	1.6471	2.77118	26.8	19.8186
TPP	5.28	11.462	7.3408	0.3473	0.2603	3.5085	0.0214	61.914	27.822	0.0978	0.0016	7.65064	2.5	1.35679
SpPSp	8.2	1.1163	7.7142	6.5895	3.5252	14.857	3.2294	3.3234	0.2653	6.25	3.6962	7.07878	5.25	3.73493
SDpSp	9.3	0.7373	0.099	0.5876	12.591	0.0016	2.8629	0.4099	0.8034	2.43	1.5097	6.8855	8.22	0.07618
SPL	8.03	4.4963	1.144	2.1584	9.0887	5.2761	3.6543	1.0677	18.529	4.75	3.3318	4.42675	23.8	2.06081
TKW	8.03	3.5032	2.9262	7.1905	3.5341	6.389	24.204	0.2938	0.6836	21.7	9.0436	6.26443	2.51	21.6351
GY	1.58	26.454	0.7187	0.7921	0.1467	19.87	6.313	1.8223	0.0363	12	1.595	7.92149	1	0.66686
HI	4.53	16.859	0.0026	0.1425	7.6304	9.425	0.0017	2.0748	3.7215	0.0776	2.253	6.56323	3.01	16.1722
BY	7.09	1.4737	17.061	2.4289	11.033	0.2411	7.1957	1.902	6.5283	3.66	4.6617	6.66837	2.89	14.8875
DB	7.56	6.6321	0.0439	0.1726	6.3096	2.8393	7.4563	19.817	9.7158	8.03	0.2149	7.90787	1.32	0.17017
DM	4.85	15.702	0.246	0.0909	10.784	1.768	5.5665	0.9104	3.1581	0.00006	25.318	6.4867	3.4	16.4363
GC	7.4	6.4621	0.485	8.5278	7.4739	10.011	0.5946	0.222	2.8981	1.76	16.851	7.45538	4.36	0.15233
SpD	8.98	0.1856	1.9039	15.093	10.418	3.2234	4.5014	0.0128	0.2563	0.434	0.0424	5.82031	14.4	2.13029
SDC	3.23	0.3523	48.869	30.655	10.008	2.7507	7.0209	0.1805	0.0212	10.1	6.0263	8.0862	0.00004	0.43976
Awnd	7.84	0.2886	10.049	24.285	1.4521	11.085	22.13	0.0025	1.345	0.00206	23.808	8.01319	0.571	0.2621
PV	69.9	20.8	8.0	1.3	49.75	28.35	12.6	8.7	0.5007	0.05816	0.0118	82.1	12.3	5.6
CV	69.9	90.7	98.7	100	49.75	78.1	90.7	99.4	99.93	99.9882	100	82.1	94.4	100
EV	10.5	3.1136	1.20668	0.1986	7.4619	4.2505	1.8936	1.3085	0.0751	0.0087	0.0018	12.321	1.838	0.84105

Remarks: SDpsp= the number of seeds per spikelet, SpPSp = the number of spikelet per spike, SPL= spike length (cm), TPP= the number of effective tillers per plant, PH= Plant height (cm), TKW= weight of thousand kernels per plant (gm). GY= Grain yield (t/ha), BY= biomass yield (t/ha), HI= harvest index, DB= days to booting, DM= days to maturity, GC= glume colour, SpDen=spike density, SDC= seed color, Awnd= awndness, EV= eigenvalue, PV= percentage of variance, and CV-= cumulative variance.

Altitudinal based PCA

Altitudinal based principal component analysis revealed that 49.73 %, 28.35 %, 12.6 %, and 8.7 % of the total variation was respectively explained by the first four components with eigenvalues greater than one, which are equal to 99.4 % of the total variation when combined (Table 3.11). Seed color, awndness, and spike density were the main contributors to the first principal component.

Species based PCA

In the case of species, the first, the second, and third principal components with eigenvalues greater than one explained 100 % of the total variation. The first and second principal components explained 82.1 % and 12.3 % of the total variation, respectively (Table 3.11).

***2.3.2.8. Cluster analysis*:** The whole genotypes were grouped into seven clusters based on the fifteen phenotypic traits studied using the ward. D method and the number of genotypes in a cluster ranged from 12 in cluster VII to 48 in cluster I (Table 3.12; Figure 3.7). The highest numbers of genotypes fall in the first cluster (48) followed by the second (47) and six (45) clusters (Table 3.12). The first cluster contained 48 genotypes which accounted for 24.49 % of the total genotypes used in the study. The second cluster contained 47 genotypes which accounted for 23.98 % of the genotypes used in the study. From the present study *Triticum durum*, *Triticum turgidum*, *Triticum aethiopicum, and only one Triticum polonicum* species were found while *triticum diccocum was absent*. The work of Getachew and Worede (1991) agrees with this study. The cluster analysis of tetraploid wheat for the entire accessions indicates the absence of clear regional grouping or there is lack of strong regional differentiation, which could be partly ascribed to gene flow. Cluster analysis indicated that landraces from the same location did not fall in the same cluster and at the same time landraces from different localities may fall in the same cluster implying that there is no relation between geographic and genetic distances. This, in turn, implies that accessions collected from the same geographic area were not necessarily closely related and different regions did not necessarily have different genetic backgrounds. The present study is in line with other similar studies (Asmamaw *et al.,* 2020; Hailu *et al*., 2006; Gashaw *et al*., 2010). The correspondence of landraces from different regions of collections might be due to exchange of seeds of the same genetic background by farmers (Gashaw *et al.,* 2010; Hailu *et al*., 2006). The clustering pattern

indicated that the accessions were distributed patchily over many clusters, showing wide variation among accessions within a particular region.

Table 3. 12 Clustering of 174 tetraploid wheat landraces along with 22 improved cultivars into seven groups using 15 phenotypic traits

Cluster	Lists of genotypes	Region					Number of genotypes	Percentage of total population (n=196)
		Shewa	Arsi	Gojam	Wello	rv		
I	5158, MCD5-4,MCD5-25,MCD6-44, 7205,7046,7581,MCD14-29,ETDW/15DZ#39,5600,ETDW/15DZ036,ETDW/15DZ063, 243703,MCD3-27, 5635, 7532, 226694, 242789, 5181, 5169, 5149, ETDW/15DZ#50,ETDW/15DZ#52, 5534, 5638, 5581, 7664,7825,242790,MCD3-14, 5618, 243706, 5913, 238891, MCD8-44,7078,MCD11-20,MCD8-2, ETDW/15DZ019,……, ETDW/15DZ#,5504,5666,5216,5728,7832,MCD12-30, 5140, 7056, 7003,7295,242793,MCD4-32,5586,	34	4	4	6		48	24.49
II	226241, MCD13-8,MCD14-29, 7063, 242784, 242781, 7641, 5434,7007,5760,7050,5182,5057,7880, DZDW17007,7626, 7060,6102,5707,5510,7629,6982,7015,5470, 7009,7242, 7673,5441,ETDW/15DZ#12,ETDW/15DZ#21……5183, 7313,7014,7020,7378,7075,MCD5-35,ETDW/15DZ004, MCD12-2,MCD12-38,MCD14-26,…….,5197,5591,MCD4-4, 7004,7568,5052,5168,MCD14-3,MCD11-28,MCD11-36,……., 7666,MCD4-32-2,7069	25	14	5	3		47	23.98
III	ETDW/15DZ#43ETDW/15DZ#068, 7647, 6975, ETDW15DZ073, ETDW/15DZ009,……. ETDW/15DZ#21ETDW/15DZ#31,ETDW/15DZ014,7507, ETDW/15DZ055,MCD3-1, Arendato, 7571, 5174, 7150, 235051,Ld-357,Malefia,5184	8		2	3	3	16	8.16
IV	5515, 6987,MCD3-19, ETDW/15DZ#34,ETDW/15DZ#35, 6971,7578,222427,ETDW15DZ012,ETDW/15DZ070, MCD4-13,MCD8-6,MCD11-18,5584,5492,7209,ETDW/15DZ#4, ETDW/15DZ#20,DZDW176,7,10,DZDW17005, DZDW17006,07,7999,5554	10	2	1	2		15	7.65
V	6914,239711,Tob66,MCD3-15,Gerardo, Quamy, Foka, ETDW/15DZ#044,Selam, Asasa,Ginchi,Boahi,5465	3	1	1		8	13	6.63

VI	DZDW17004,DZDW1702,04,5729,MCD10-11, 243717, 5373, 5369, 5576,7827,5172,7195,242783, 5342,6968,7798, 242786, ETDW/15DZ043,DZDW17010.8,MCD10-3,MCD10-16, MCD10-44, 5609, 5214, 6955, 6933, 5171, 5491, 5142, 7477, 7713,7823,7076,5044,243700,7464,5893, MCD11-7, MCD11-25,MCD13-5,MCD14-19,7826,MCD7-35, 7762, 7690, 242792, 242782, 243701, 5166,214370,MCD5-18,7683,7191	26	4	10	5		45	22.96
VII	Werer,DZDW17002,Yerer,Alemtena,Tesfaye,Ejersa,Mangudo,Fetan ,DZ2018,Mukiye,Bekelcha,Cocorit/71	1				11	12	6.12
	Total	105	28	24	19	22	196	

NB: numbers or Codes refer to the genotypes used in the study which were detailed in the lists of the materials used in the experiment

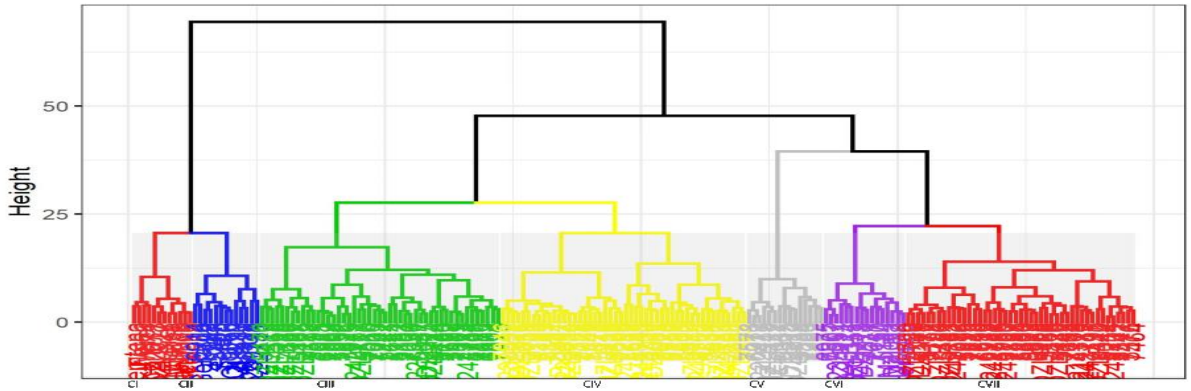


Figure 3. 7 Dendrogram showing the clustering pattern of the 174 tetraploid wheat landraces along with 22 improved cultivars based on the fifteen phenotypic traits

The intra- and inter-cluster distances were obtained based on the average group linkage produced from the centroid (Table 3.13). The minimum inter-cluster distance (2.14) was recorded between clusters I and VI while the maximum inter-cluster distance (5.58) was obtained between clusters I and VII. The minimum (2.78) and the maximum (3.4) intra-cluster distance were recorded for clusters III and V, respectively.

Table 3. 13 Intra- and inter-cluster distances of tetraploid wheat genotypes based on average group linkage produced from centroid using 15 phenotypic traits

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	2.96						
Cluster2	2.20	3.09					
Cluster3	3.50	3.62	2.78				
Cluster4	4.19	4.67	4.48	2.99			
Cluster5	4.34	3.85	5.06	6.04	3.40		
Cluster6	2.14	2.77	2.78	3.94	4.70	2.84	
Cluster7	5.58	5.21	5.00	6.37	3.86	5.27	2.80

Regional clustering revealed that the landraces from Shewa were grouped with the landraces from Gojam and Wello while the landraces from Arsi remained solitary (Figure 3.8). This result also agrees with biplot of PCA of regions of collections in figure 3.6.

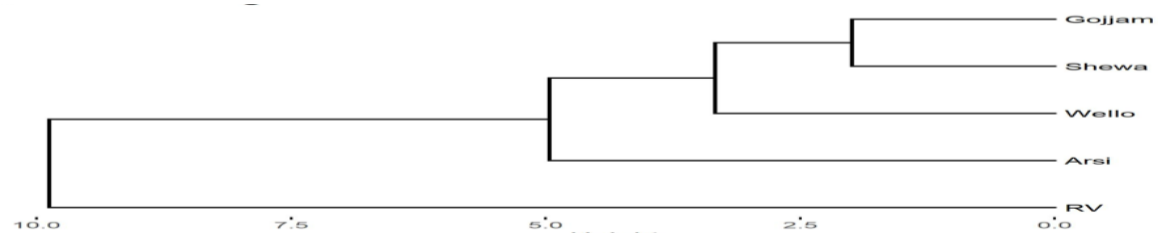


Figure 3. 8 Dendrogram of the regions of collections of 174 landraces along with 22 improved varieties based on 15 phenotypic traits. RV= Released varieties

The released varieties were separately clustered. The maximum Euclidean distance (8.21) was observed between the landraces from Wello and released varieties followed by the distance (7.64) between the landraces from Arsi and the released varieties. The minimum Euclidean distance (1.99) was recorded between the landraces from Shewa and Gojam followed by the distance (2.38) between the landraces from Shewa and Wello (Table 3.14). In line with to this study, the work of Hailu *et al.* (2010) indicated that released varieties stood solitary; however, contrary to this study materials from Gojam, Wello, Shewa, and Gonder clustered together and Arsi and Bale to one cluster.

Table 3. 14 Euclidean distance between regions of collections of 174 tetraploid Wheat landraces and 22 improved cultivars based on 15 phenotypic traits

	Wello	Shewa	Arsi	Gojam	Released Varieties
Wello	0				
Shewa	2.38	0			
Arsi	4.95	3.63	0		
Gojam	3.62	1.99	4.01	0	
Released Varieties	8.21	7.17	7.64	6.81	0

From the present study, *Triticum durum, Triticum turgidum, Triticum aethiopicum, and only one Triticum polonicum* were found. Species clustering based on the means of the 15 phenotypic traits indicated that *Triticum turgidum* and *Triticum aethiopicum* clustered together while *Triticum durum* and *Triticum polonicum* remained solitary (Figure 3.9). Similar to this study, Hailu *et al*. (2006) reported that *Triticum turgidum and Triticum aethiopicum* were clustered together in tetraploid wheat of Ethiopia. Opposite to this finding, Hailu *et al.,* (2010) *Triticum durum* and *Triticum aethiopicum* clustered together.

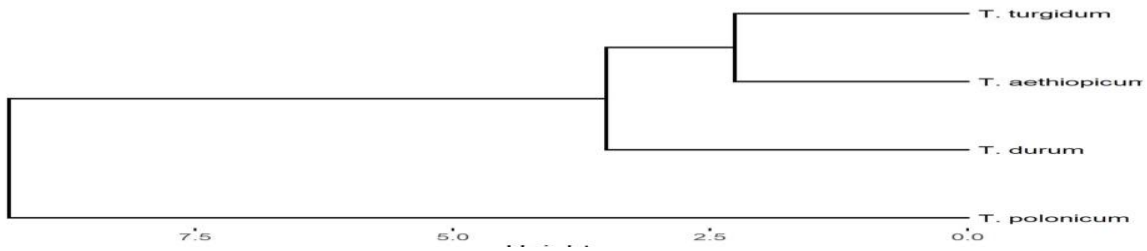


Figure 3. 9 Dendrogram of species of 174 landraces along with 22 improved varieties based on 15 phenotypic traits

Table 3. 15 Euclidean distance between species of 174 tetraploid wheat landraces along with 22 improved cultivars based on 15 phenotypic traits

Species	*T.turgidum*	*T.durum*	*T.aethiopicum*	*T.polonicum*
T.turgidum	0			
T.durum	2.80	0		
T.aethiopicum	2.17	3.22	0	
T.polonicum	7.15	7.69	6.85	0

Altitudinal clustering revealed that altitude classes 8 (≥2921), 4 (2201-2380), 6 (2561-2740), and 5 (2381-2560) were grouped together (Figure 3.10). On the other hand, landraces from altitude class 7 (2741-2920) and 3 (2021-2200) were grouped together. The landraces from class 1 (≤1840) and class 2 (1841-2020), obtained from low altitudes in Arsi, Bale and East Shewa, remained solitary. The highest Euclidean distance was 8.75

followed by 7.58 between altitudinal classes 1 (≤1840) and 2 (1841-2020) and between altitudinal classes 2 (1841-2020) and 4 (2201-2380). The minimum Euclidean distance (1.81) was between altitudinal classes 5 (2381-2560) and 6 (2561-2740) (Table 3.16).

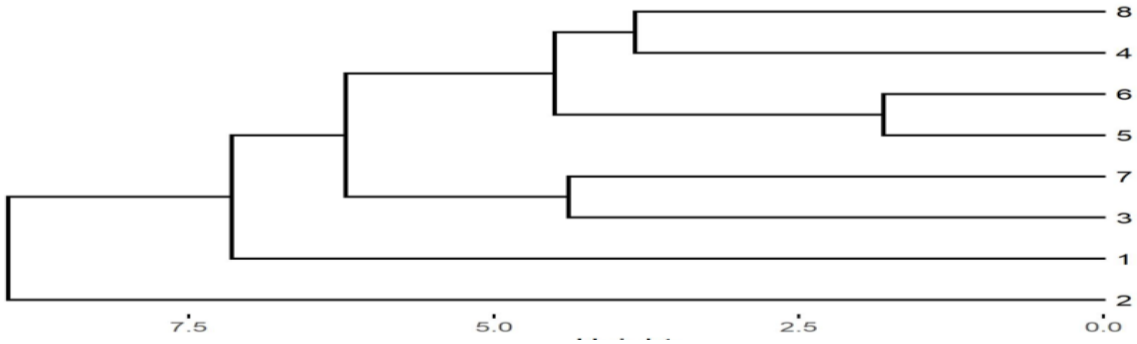


Figure 3. 10 Dendrogram of altitudinal classes of 174 landraces based on 15 phenotypic traits

Table 3. 16 Euclidean distances between altitudinal classes of 174 tetraploid Wheat landraces based on 15 phenotypic traits

Altitude class	Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8
Class1	0							
Class2	8.75	0						
Class3	6.77	6.48	0					
Class4	5.79	7.58	5.31	0				
Class5	5.99	6.96	3.99	3.93	0			
Class6	5.52	7.03	4.58	3.85	1.81	0		
Class7	6.09	6.86	4.39	5.74	4.66	4.50	0	
Class8	5.23	6.23	3.86	3.84	3.22	3.64	4.93	0

3.5. Conclusion

There is high phenotypic diversity in tetraploid wheat of Ethiopia regarding glume color, seed color, and spike density, which is in agreement with the report of (Vavilov 1929) and (Jack R. Harlan 1969) implying that Ethiopia is a centre of diversity of for the crop; however, low variability was observed in their awndness. Intermediate mean diversity was observed across regions of collections (0.57±0.07), altitudinal class (0.49±0.05), and species level (0.57±0.08) for the four qualitative traits. There is a pattern of variation across altitudinal class for the four qualitative traits. The present study result provided preliminary indications that tetraploid wheat of Ethiopia holds huge genetic variation, which could be used as potential input in the breeding programme. The tetraploid wheat genotypes studied showed highly heritable variation accompanied with high expected genetic advance for the number of seeds per spikelet implying that selection may be effective for this specific trait. Thus, the top 5% of the tetraploid wheat genotypes for the number seeds per spikelet including DZDW17002, Alemtena, Fetan, Malefia, DZ2018, Tob66, Werer, Selam, Cocorit/71, and Ejersa could be used as potential parents for crossing to develop improved varieties for this trait. Most of the traits studied showed positive association with grain yield implying that improving either one or more of these traits could result in high grain yield. Further work is needed to evaluate these and additional traits under different environmental conditions to assess their suitability for contributing to future breeding programmes. The grouping of accessions by multivariate methods would be of practical value to wheat breeders in that representative accessions may be chosen from different clusters for breeding purposes and to define strategies for further collection of the germplasm. Moreover, future germplasm collections should aim to cover collections from as many geographically and climatically different areas as possible, rather than making collections within the region, to enable the country to conserve the huge genetic diversity.

CHAPTER IV

4. Genetic Diversity in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia based on grain quality traits

Abstract

*The need for high-quality wheat-based products is growing. A total of 174 landraces along with 22 released varieties of tetraploid wheat (Triticum turgidum spp) germplams were evaluated for their genetic diversity based on grain quality traits at Sinana and Debrezeit during 2020 in simple lattice with two replications. Combined analysis of variance revealed highly significant variation ($p<0.001$) among the genotypes and for genotype by location interaction for the traits studied. The genotypes showed a wide range in their thousand kernel weight, zeleny index, and vitreousness percentage. Grain yield and protein content, respectively, showed high and low value of PCV and GCV; however, an intermediate value of PCV and GCV was observed for gluten content and thousand kernel weight. Zeleny index and vitreousness percentage showed high PCV and intermediate GCV; however, protein showed the lowest values of PCV and GCV. The estimate of broad sense heritability (H^2) ranged from 33.57% for vitreousness percentage to 66.36% for zeleny index. The values of genetic advance and GAM observed were respectively high for grain yield (31.94, 26.91) and zeleny index (31.89, 29.34); intermediate for thousand kernel weight (18.27, 16.59), gluten content (15.70, 13.55), and vitreousness percentage (15.75, 12.97) and low for protein content (9.81, 8.44). Grain yield showed a highly significant positive correlation with thousand kernel weight (0.43***) and vitreousness percentage (0.19**) and a positive correlation with protein content (0.08), zeleny index (0.06) and gluten content (0.1). Cluster analysis grouped the whole genotypes into six groups. The first three principal components explained 69.11%, 89.99 %, and 95.06 % of the total variation in the entire genotypes, regions of collections, and altitudinal classes respectively. In the entire genotypes, vitreousness percentage, zeleny index, and gluten content; protein content, gluten content, and TKW and grain yield and TKW respectively contributed to the variation in the first, second, and third principal components. Eleven landraces and eleven improved varieties were selected for production based on their grain quality traits performance.*

Key words: Genetic diversity; Grain quality traits; Heritability; Tetraploid wheat.

4.1. Introduction

There is an increasing need for high-quality foods, particularly wheat-based products (Bhalla *et al.*, 2017). With the increase in grain production, research is giving special attention to improving wheat's technological qualities such as the strength of flour, gluten quantity and quality, protein content, and others attributes (Keler *et al.*, 2020).

Wheat grain is an excellent staple food with numerous nutritional and health-beneficial compounds due to the special bread that is only made from its flour (Kong *et al*., 2013) that make wheat unique among cereal grains. The insoluble storage proteins of wheat collectively called 'gluten,' give wheat its functional properties (Tipples *et al.*, 1994). Depending on the unique properties of the protein it contains, wheat grain has many foods and non-food benefits (Shewry 2009) making the grain quality of wheat the second crop yield (Keler *et al.*, 2020).

In the past, grain quality was not a decisive criterion for variety selection. The breeding programs gave emphasis only to the improvement of grain yield and disease resistance (Drikvand *et al*., 2018). Nowadays, however, the grain quality of wheat is becoming central to food processing and nutritional value and is a decisive factor for consumer acceptance and commercial value of wheat cultivars (Wang *et al*., 2017) due to the current increase in the requirement for high-quality food, especially for wheat-based products (Bhalla and Singh 2017). Things have changed due to changing food habits, increasing urbanization, and trends toward raising middle-class society (Wang *et al*., 2017) that make grain quality of wheat central to human well-being (Kong *et al*., 2013). Thus, in addition to enhancing yield, improving grain quality of wheat is also vital. As a result, apart from grain yield and disease resistance, quality is becoming a solid criterion of wheat variety selection (Johnson and Johnson 2017) and improving wheat grain quality. Hence, critically looking for better quality varieties suitable for preparing different end products is becoming a criterion for developing new varieties (Peña *et al*., 2002).

To meet the current increase in the requirement for high quality wheat-based food products (Bhalla *et al.*, 2017), assessing the diversity of the existing landraces for their grain quality traits is crucial and a prerequisite for plant breeding and genetic resource conservation programmes (Laidò *et al*., 2013). Genetic diversity is a critical input for successful adaptation and sustained improvement in crops (Marone *et al*., 2012). Knowledge of genetic diversity is very important to identify, conserve, and a prerequisite to efficiently utilize useful materials from the existing landraces (Ruiz *et al*., 2012), is the basis for

parent selection in crossing program, and has significant impact on improvement of crops (Zhang *et al*., 2011). A better knowledge of the collected material can be obtained using morphological, biochemical, and molecular markers (Teklu and Hammer 2008), which can help us in the assessment of genetic variability among genotypes within landraces, populations, and species (Eticha *et al*., 2005). Ethiopia is a center of diversity for tetraploid wheat landraces (Haile *et al*., 2013; VAVILOV, 1929). These landraces are untapped primary sources of desirable genes useful for economically important traits such as grain quality. They are endowed with ample genetic resources vital to meet the diversified goal of breeding (Haile *et al*., 2013; Muleta *et al*., 2017). Despite the existing potential, Ethiopia is importing wheat from abroad due to huge gap between production and consumption (Abate 2018; Brasesco *et al*., 2019) to fill gap for industry (Haile *et al*., 2013). This necessitates the development of wheat varieties that are high yielding with the required grain quality traits through the continuous assessment of the landraces. This requires a continuous supply of new germplasm as a source of desirable genes (Asmamaw *et al*., 2019) for which tetraploid wheat are the potential source.

The study of Haile *et al*. (2013) and that of Teklu and Hammer (2008) focused mainly on phenotypic diversity. However, phenotypic diversity alone cannot clearly indicate the existing genetic diversity in the landraces. Alemu *et al*. (2020) made genome wide association mapping for grain shape and color traits in Ethiopian durum wheat (*Triticum turgidum* ssp. durum). Grain shape and color strongly influence yield and quality of durum wheat (Alemu *et al*., 2020). The existing landraces collected and maintained by Ethiopian Biodiversity Institute were not fully studied to effectively exploit the existing germplasm for different economically important traits. Therefore, this study was aimed at assessing the extent and pattern of diversity of tetraploid wheat germplasms of Ethiopia based on grain quality traits.

4.2. Material and Methods

4.2.1. Planting material

Seed samples from 196 genotypes including 174 landraces along with 22 varieties that were evaluated at Sinana and Debrezeit Agricultural Research Center in simple lattice design with two replications were assessed for their grain quality traits at Sinana grain quality laboratory.

4.2.2. Grain quality traits data collection and analysis

Grain protein content (%), moisture content (%), and gluten content (%) were determined from the seed samples of each location harvested from the two replications using MINIFRA SmarT grain analyser (Mininfra SmarT Grain Analyser Operating Manual, 2013). Vitreousness percentage was estimated from the seed sample using illuminated purity work board model S-115 by using transmitted light according to ICC standard number 129 (ICC, 2000). Moreover, grain yield (t/ha) and thousand kernel weight (gm) were also recorded.

Grain yield per plot: the grain weight of all plants grown per plot taken from each genotype, moisture content adjusted to 12.5%.

Thousand kernel weight (gm): the weight of thousand grains of wheat genotypes. The weight of 1000 wheat kernels for different varieties ranges from low (30-35g), medium (36-45 g) to high (≥ 45 g).

4.2.3. Grain quality traits data analysis: For data analysis, landraces from regions with small sample sizes were included into adjacent regions to reduce experimental errors due to small sample sizes. In doing so, landrace from Jimma along with, one landrace with unknown location, were included into Shewa, five landraces from Gonder and two from Agaw Awi were included into Gojam, two landraces from Tigray into Wello and three landraces from Bale were included into Arsi.

The data were standardized to a mean of zero and a variance of unity to avoid differences in scales used for recording data on the different traits (Sneath and Sokal 1973) before undertaking a series of multivariate analyses using the appropriate procedures of R software (version 4.1.1) (Harris 2018).

Hierarchal clustering of the genotypes was performed for the entire genotypes, regionally, attitudinally, and at the species level using R software version 4.1.1 for the standardized data to mean zero and unity variance (Sneath and Sokal 1973). Analysis of variance was carried out using R software (version 4.1.1) (Harris 2018). The combined ANOVA was performed by the model: Y= Replication + genotypes + Location + Genotypes x Location +Replication: block.

The variability of each quantitative grain quality trait was estimated using simple statistical measures such as mean, range, phenotypic and genotypic variances, and coefficient of variation. The phenotypic, genotypic, environmental, and genotype by environment interaction variation and coefficient of variations, broad sense heritability (H^2) and genetic advance were computed using excel following the formula suggested by Singh and Chaudhary (1985) and Allard (1960).given below:

Genotypic variance ($\delta^2 g$)

$$\delta 2g = (MSg - MSgl)/rl$$

where MSg stands for mean square of genotype, MS_{gl} represents mean square due to genotype by environment interaction, l is number of locations and r stands for number of replications.

Genotype by environment interaction variance (δ^2_{gl})

$$\boldsymbol{\delta gl^2} = (MS_{gl} - MS_e)/r$$

where MS_{gl} represents mean square due to genotype by environment interaction, and MSe represents combined error mean square.

Phenotypic variance (δ^2_p)

$$\delta p^2 = \delta g^2 + \left(\frac{\delta gl^2}{l}\right) + \left(\frac{\delta e^2}{rl}\right)$$

Estimates of coefficients of variation were obtained as follows.

Phenotypic coefficient of variation (PCV)

$$PCV = \frac{\sqrt{\delta p^2}}{\mu} x100$$

where PCV = phenotypic coefficient of variation, δ^2_p = phenotypic variance and μ = population mean for the trait considered.

Genotypic coefficient of variation (GCV)

$$GCV = \frac{\sqrt{\delta g^2}}{\mu} x100$$

where GCV = genotypic coefficient of variation, δ2g = genotypic variance and μ = population mean for the trait considered.

Environmental coefficient of variations (ECV)

$$ECV = \frac{\sqrt{\delta e^2}}{\mu} x100$$

Genotype by environment interaction coefficient of variation (GECV)

$$GECV = \frac{\sqrt{\delta gl^2}}{\mu} x100$$

where, δ^2_{gl} = genotypic by environment interaction variance and $\boldsymbol{\mu}$= population mean for the trait considered.

PCV and GCV values below 10%, 10%-20% and above 20% were low, intermediate and high respectively (Khorgade, et al. 1985).

Broad sense heritability (H^2) and genetic advance

Heritability in broad sense was estimated based on the formula given by Allard (1960). Heritability for the two locations was calculated using the formula:

$H^2 = (\delta 2g/\delta 2p)\ x\ 100$, where $\delta 2p = \delta 2g + \left(\frac{\delta 2gl}{l}\right) + \left(\frac{\delta 2e}{rl}\right)$. Broad sense heritability was categorized as low (below 40%), medium (40%-80%), and high or very high (above 80%).

Expected genetic advance under selection was calculated following Allard (1960), at 5% selection intensity, as: $GA = (K)\ (\delta p)\ (H^2)$, where GA = expected genetic advance, K represents a selection differential that varies based on the selection intensity and is equal to 2.056 if one chooses 5% of the genotypes, δ_p stands for phenotypic standard deviation and H^2 represents broad sense heritability. Genetic advance as percentage of the mean was calculated as$GA\ (\%\ of\ mean) = (\frac{GA}{\mu})\ x100\%$, where GA = genetic advance and μ = population mean for the trait considered. Genetic advance as percentage of mean was categorized as suggested by (Johnson *et al*., 1955) as low (<10%), moderate (10-20%), and high (>20%).

Cluster analysis: Hierarchal clustering for the entire data set, regions of collections, altitudinal classes, and at species level was performed based on ward.D linkage method with Euclidian distance, obtained by dist () function, using R software version 4.1.1(Harris 2018). The distances between clusters and within clusters were calculated using the average linkage method of Euclidean distance based on their centroids.

Principal component analysis: Principal component analysis was performed for the entire data set, regions of collections, altitudinal classes, and at species level to identify the most important traits contributing to the total variations observed among the accessions, regions of collections, and altitude classes using princomp () function in R software version 4.1.1 (Harris 2018**).** As suggested by Johnson (2007), principal components with Eigen values greater than one were considered.

Correlation coefficient analysis: Pearson correlation analysis was performed for quantitative traits using R software (version 4.1.1) (Harris 2018**).**

Comparison of selected genotypes with the original population: to compare with the original population, the means of the top 5% genotypes for each trait were independently computed. The student t-test table was used to determine the significance of the difference between the sample mean and population parameter. When the calculated t-value is higher than the tabulated t-value, the difference is deemed significant. To compare the performance of the 5% best-selected genotypes with the size of a population, the absolute t-value was obtained using the student's t-test formula as follows:

$$T = (\bar{x} - \mu)/(S/\sqrt{n})$$

Where n is the number of genotypes selected from the size of a population for better performance, $\bar{x}$ is the mean of the genotypes that were selected, S is the sample standard deviation, and is the mean of the size of the population.

4.3. Results and Discussion

4.3.1. Analysis of variance

Homogeneity of the error MS from individual analysis of variance at Sinana and Debrezeit was checked following the formula of Gomez and Gomez (1984) described as: F=Larger error MS/smaller error MS. The error variances were homogeneous for all the triats studied for diversity analysis. Moisture content was adjusted to 12.5 percent for gluten content, protein content, zeleny index, vitreousness percentage, and grain yield before analysis was conducted. Combined analaysis of variance revealed that there was highly significant variation ($p< 0.001$) among the genotypes for all grain yield, thousand kernel weight, gluten content, protein content, zeleny index, and vitreousness percentage (Table 4.1) suggesting that a substantial genetic diversity in the population for the grain quality traits studied. This is crucial for breeding programs providing a broad base for selecting superior genortypes. Except vitreousness percentage, all the traits studied were highly signficantly affected by the location. Moreover, grain yield, gluten content, protein content, zeleny index, and vitreousness percentage were highly significantly affected by the genotype by location interaction ($P<0.001$) implying test site had a signficant impact in these traits (Studnicki *et al*., 2016) and that the performance of genotypes varies across different environments might be because of weather, farming practices, soil charcateristics or field management, which affect how genes are exprerssed (Yao *et al*., 2008; Persaud *et al*., 2022); thus, highlighting the importance of considering environmental conditions in breeding decisions. Except vitreousness percentage, all the traits studied were highly significantly affected by the site of study. The significant variation among genotypes suggests that there is a substantial genetic diversity in the population for the grain quality traits studied. This diversity is crucial for breeding programs, as it provides a broad base for selecting superior genotypes. The significant variation among genotypes and the substantial GxE interaction underscore the need for comprehensive breeding strategies that account for both genetic potential and environmental influences. This dual approach can lead to the development of robust and high-quality grain varieties that perform well across diverse conditions.

Amiri *et al.* (2018) also reported that there was significance difference between genotypes, significant effect of environment on thousand kernel weight, grain yield and grain protein content and significant effect of genotype by environment interaction on protein content

and thousand kernel weight in bread wheat genotypes. Similarly, significant variation between genotypes in grain protein content was reported by Mesele *et al.*(2015). Similar to this study, Taneva *et al.*(2019) reported highly significant variation among genotypes, across locations and in the effect of genotype by environment interaction in protein content, vitreousness percentage, thousand kernel weight, and grain yield of durum wheat except for the significant effect of location in vitreousness percentage.

Table 4. 1 Mean squares of grain quality traits from combined ANOVA of 196 tetraploid wheat genotypes from Ethiopia

SV	df	TKW	GY	Gluten	Protein	ZI	VITP
Replication	1	627***	5.21*	162***	17.8***	617**	340
Genotypes	195	106.94***	5.82***	65.72***	6.01***	377.56***	976***
Location	1	10489***	188.54***	3913***	1257.5***	17056***	633
Genotypes: Loc	195	44***	3.39***	32.93***	2.9***	127***	648.4***
Rep:block	26	40*	1.59	25**	1.9**	140**	349.1*
Pooled error	365	21.2513	1.03539	10.4518	0.79971	58.6247	185.768
	mean	34.24	3.25	26.49	13.3	41.57	68.44
	CV	14.01	33.1	13.08	7.13	19.84	20.52
	lsd	6.67	1.5	4.82	1.32	11.47	19.53

NB: SV = sources of variation, df = degrees of freedom, CV= coefficient of variation, lsd= lest significant difference, Loc= location, GY = grain yield, TKW = thousand kernel weight, ZI = zeleny index, and VITP = vitreousness percentage.

4.3.2. Performance and patterns of variation of grain quality traits of the genotypes: The coefficient of variation revealed high variability among the genotypes with respect to grain yield (t/ha) and vitreousness percentage; intermediate variability for thousand kernel weight, gluten content, and zeleny index and low variability for moisture content and protein content (Table 4.1, Annex 3).

The genotypes showed high range values for vitreousness percentage (69.7), zeleny index (43.71), and thousand kernel weight (21.3); intermediate for gluten content (17.37) and low for protein content (4.76) and grain yield (4.65) (Table 4.2). The patterns of quantitative traits variation also revealed that there was high variation among the tetraploid wheat of Ethiopia. Thousand kernel weight was on average 34.24 gm ranging from 23.6 gm to 44.9 gm, indicating high range values (21.3gm). Sixty-three of the genotypes, including sixteen improved varieties and fourty seven landraces, had medium thousand kernel weight (36-45), 105 genotypes were in the low kernel weight group (30-35), and 28 genotypes showed very low kernel weight (<30 gm). The mean and range values for thousand kernel weight, grain yield, and protein content reported by Taneva *et al*., (2019) is similar to this finding.

Similar result was also reported by Mesele *et al*. (2015) on a range (11.93% to 15.43%) and mean (13.79%) values of grain protein content in bread wheat genotypes. Amiri *et al.* (2018) also reported almost similar result on minimum (24.10, 11.53), maximum (49.14, 12.65), and mean (39.01 ±0.62, 12.01 ± 0.03) values respectively for thousand kernel weight and grain protein content of bread wheat under rain-fed conditions. A similar result on zeleny index mean (45.15ml) and range (21.10 to 56.83) in wheat genotypes was reported by Modoran *et al.* (2010). Yalcin and Maden (2020) also report similar result on zeleny index range values (25.5 ml to 65 ml) in wheat flour. Similar result was reported by Kumar (2020) on protein content mean (13.01) and range (12-14) on wheat genotypes under restricted irrigation conditions. Sayaslan (2012) also reported 10.7 % to 16.8 % range values of protein contents in durum wheat landraces.

Table 4. 2 Range, mean, minimum, and maximum of grain quality traits of the 174 landraces along with 22 improved varieties

		Grain quality traits					
Category		TKW	GY	Gluten	Protein	ZI	VITP
Whole genotypes	Mean	34.24±0.28	3.25±0.06	26.49±0.21	13.3±0.06	41.57±0.57	68.44±1.12
	Min	23.6	1.33	17.89	11.14	19.48	28.48
	Max	44.9	5.98	35.26	15.9	63.19	98.18
	Range	21.3	4.65	17.37	4.76	43.71	69.7

Grain yield was on average 3.25 (t/ha) with a minimum and maximum average values of 1.33 t/ha and 5.98 t/ha, respectively, and a range of 4.65 t/ha. Four landraces (239711, 7060, 5638, and 242789) and two improved varieties (Tob66 and Foka) gave 5.98, 4.99, 4.75, 4.72, 4.7, and 4.67 average grain yield (t/ha), respectively and these landraces, except the landrace 242789, were in acceptable thousand kernel weight; that is, 7060, 5638, and 239711 recorded medium mean values of TKW while 242789 is low in its TKW.

The average gluten content was 26.49 % with a minimum and maximum average value of 17.89 % and 35.26%, respectively, and an intermediate range value (17.37). The minimum gluten content (17.89%) was recorded from the landrace 242792 at an altitude of 2330 m.a.s.l. in East Shewa while the maximum gluten content (35.26%) was recorded from the landrace 238891 obtained from Bale at an altitude of 1900 m.a.s.l. Twenty seven landraces gave gluten content, which is higher than the mean value of gluten content recorded by Ejersa variety (26.31 %) that is top performing among the improved varieties (Annex 3). The gluten content recorded in the present study ranged from 17.89% to 35.26%. The

result is in the range of gluten content, which can vary from 11 % to 59 % depending on different factors such as variety, location, heat, and moisture (Keler *et al.*, 2020).

Protein content was on average 13.3 % with a minimum and maximum average value of 11.14 % and 15.90%, respectively and a range of 4.76%, which is low. The minimum protein content (11.14%) was recorded from the landrace 242792 at an altitude of 2330 m.a.s.l from East Shewa while the maximum protein content (15.90%) was recorded from landrace 6955 obtained from Arsi at an altitude of 2440 m.a.s.l. All the genotypes gave above 11% protein, which is suitable for bread-making (Bushuk and Rasper 1994). A total of 42 genotypes, including forty landraces along with two improved varieties (Ld-357 and Coccorit/71), gave protein content higher than 14%, which is demanded by the modern pasta manufacturing industry. The protein level for quality pasta products must be between 12% and 16 % (at 14% mb) (Petrova 2007). On the other hand, 115 (59%) of the genotypes gave the mean value of protein content higher than 13 percent, excellent pasta cooking quality (Royo *et al*., 2005), which is a standard in quality throughout the grain industry (Mekonnen *et al*., 2020) and a key quality parameter in durum that is important to obtain a good texture of cooked pasta (Kadkol and Sissons 2016). All the improved varieties along with 171 landraces gave higher than 11.50% protein content, which is a seed protein level in the hard wheat (Varzakas *et al.*, 2014). Four landraces (6955, 5728, 5504, and 5465) along with the improved variety Coccorit/71 gave the mean value of protein greater than 15% and among these landraces, landraces 5728 (4.52 t/ha, 36.5 gm) and 5465 (4.45 t/ha, 37.81 gm) gave medium thousand kernel weight and high grain yield (t/ha).

Zeleny index was on average 41.57 ml with a minimum and maximum value of 19.48 ml to 63.19 ml and a high range value (43.71). The maximum zeleny index (63.19 ml) was recorded from the landrace 5184 from East Shewa at an altitude of 2300 m.a.s.l. The minimum zeleny index (19.48 ml) was recorded from the landrace 5142 obtained from North Shewa of Oromia at an altitude of 2893 m.a.s.l. Zeleny index ranges from 0 to 80 and wheat having a zeleny Index below 20 is generally regarded as unsuitable for baking (AGTEC-Org 2011). Except the landrace 5142, all the genotypes had average zeleny index above 20 ml.

Vitreousness percentage was on average 68.44 % with a minimum and maximum value of 28.48 % and 98.18 %, respectively, and a high value of range (69.7). The minimum vitreousness percentage (28.48%) was recorded from the landrace 6955 obtained from Arsi at an altitude of 2440 m.a.s.l. The maximum vitreousness percentage (98.18) was recorded

from the landrace 214370 from East Shewa at an altitude of 1975 m.a.s.l. The landrace 214370, which had highest vitreousness percentage had a low yield (2.21 t/ha) and was not included in the selected genotypes based on their performance. Vitreousness is one of the international key quality-grading factors for durum wheat and the best milling quality will be found in those wheats of highly vitreous kernels (Petrova 2007). Vitreousness in wheat is regarded as high (above 75 %), medium (between 60 % and 75 %), and low (below 60 %) (Korkut *et al*., 2007). In the present study, 77, 63, and 56 of the genotypes had vitreousness value above 75 %, between 60 % and 75 %, and below 60 %. A total of 51 genotypes, including sixteen improved varieties along with 35 landraces, recorded more than 80% vitreous percentage, which is the acceptable minimum value (Petrova 2007) and these landraces are especially suited to the production of pasta because of their highly vitreous grain or high milling yield of semolina (Bushuk 1998) and the remaining genotypes recorded vitreousness percentage less than 80%, which is undesirable for pasta milling and non-vitreous wheat kernels produce a higher yield of flour, a low value by-product, in comparison to semolina during milling (Neethirajan *et al*., 2006).

Table 4. 3 Selected tetraploid wheat genotypes based on grain quality traits performance

GEN	VITP	TKW	GY	SC	GC	MC	PC	ZI	BY
7060	92.03	37.3	4.99	2	31.06	10.38	13.67	49.65	17.5
Tob66	82.6	41.5	4.7	3	28.35	10.42	13.15	47	13.13
Foka	89.96	41.75	4.68	3	27.18	10.48	12.97	49.47	10
6914	81.81	41.55	4.6	3	28.31	10.4	12.82	43.72	14.38
Yerer	88.85	40.3	4.54	3	29.07	10.4	13.69	55.64	9.38
ETDW/15DZ#044	87.51	44.65	4.45	1	23.17	10.28	11.87	47.67	14.38
5510	85.61	39.35	4.28	3	28.86	10.4	13.47	47.29	11.25
Quamy	85.98	41.9	4.27	3	27.61	10.6	13.21	45.61	13.75
Mangudo	83.17	41.1	4.15	3	27.82	10.26	12.92	53.95	10
Ginchi	86.21	41	4.04	3	24.89	10.48	11.97	50.6	7.5
Asasa	93.35	39.4	4.01	3	25.45	10.33	12.93	53.87	11.25
7209	88.51	37.9	3.88	2	24.48	10.58	12.26	43.74	15.63
7666	81.69	38.25	3.84	3	24.04	10.52	11.91	43.18	12.5
Boahi	88.29	39.78	3.79	3	27.17	10.6	13.2	53.12	10
226694	80.16	36.2	3.77	2	30.07	10.6	13.91	48.31	12.5
5893	80.99	39.05	3.77	1	23.92	10.45	12.79	47.49	10
Bekelcha	80.69	36.25	3.69	3	26.37	10.3	12.79	47.53	9.38
Selam	89.1	43.3	3.61	3	26.85	10.41	12.98	54.63	14.38
DZDW17002	91.32	36	3.59	3	30.26	10.38	13.95	57.61	7.5
6975	82.87	36.95	3.57	1	25.79	10.43	12.59	44.41	12.5
DZDW17007	91.66	37.35	3.4	3	31.99	10.58	14.49	54.64	11.88
Werer	91.75	39.35	3.3	1	28.27	10.34	13.34	56.53	13.13

Remarks: VITP= vitreousness percentage, TKW= thousand kernel weight, GY= grain yield, SC= seed color, GC= Glum color, MC= moisture content, PC= protein content, ZI= Zeleny index, and BY=Biomass yield.

A total of twenty two genotypes, including eleven improved varieties (Tob66, Foka, Yerer, Quamy, Mangudo, Ginchi, Asasa, Boahi, Bekelcha, Selam, and Werer) and eleven landraces (7060, 6914, ETDW/15DZ#044, 5510, 7209, 7666, 226694, 5893,DZW17002,6975, and DZW17007), were found to be top performing based on their grain quality traits: medium thousand kernel weight, high average grain yield, and acceptable level of protein content, gluten content, vitreous percentage, and zeleny index (Table 4.3). Except landraces 7060 and 226694, which had red seed color, other selected landraces and improved varieties had either white or purple (yellow) seed color, which is most preferred for pasta production. On the other hand, four landraces (7060, 7209, 6914, and ETDW/15DZ#044) and four improved varieties (Selam, Quamy, Tob66, and Werer) were the potential genotypes in areas where feed source is scarce due to their high average biomass yield in addition to their grain quality traits (Table 4.3).

4.3.3. Comparison of selected genotypes with the original population

It was evident that most genotypes of tetraploid wheat showed superior grain quality traits performances when compared to the population mean and the released varieties' mean (Figure 4.1). Most of the genotypes outperformed population mean with respect to vitreousness percentage and gluten content while they outperformed the mean of released varieties with respect to gluten content and protein content. The numbers of genotypes outperforming the population mean were higher than the numbers of genotypes outperforming the mean of released varieties with respect to all grain quality traits except gluten content and protein content. Overall, there were several genotypes showing superior performance over the population mean; however, most genotypes showed superior performance over the mean of released varieties in their gluten content and vitreousness percentage.

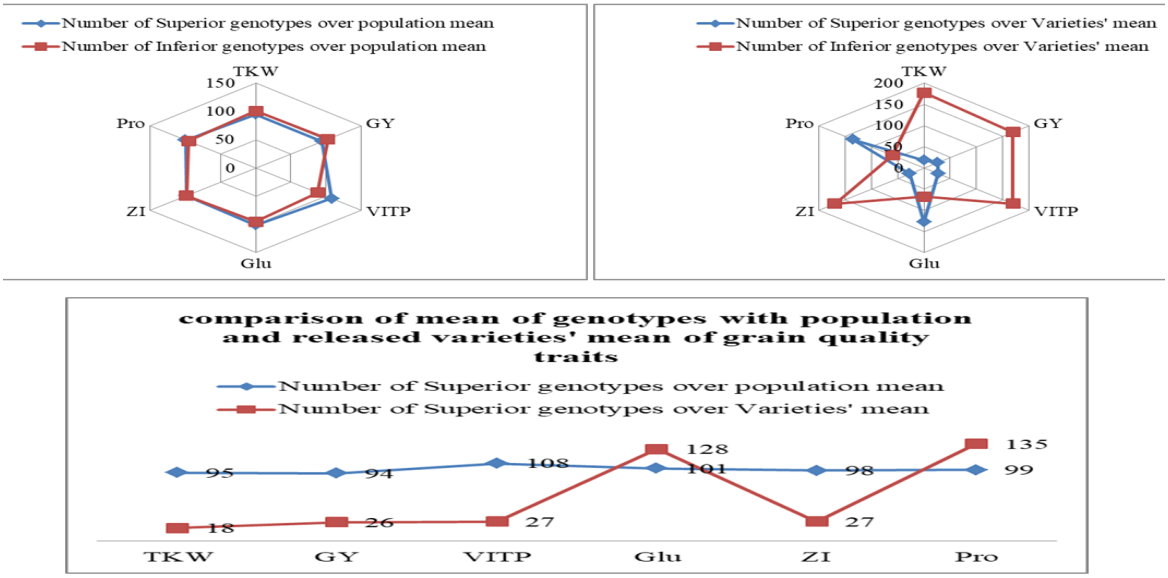


Figure 4. 1 The number of superior and inferior tetraploid wheat genotypes over (A) the population means and (B) the mean of released varieties grown during 2020 at Sinana and Debrezeit based on grain quality traits. Remark: Glu= gluten content, Pro = protein content, GY = grain yield (t/ha), TKW= thousand kernel weight (g), ZI = zeleny index, and VITP = vitreousness percentage.

The performance of the top 5% of tetraploid wheat genotypes is compared with the population means and the mean of the released varieties (Table 4.4) based on their grain quality traits. Many other superior genotypes for grain quality traits of agronomic significance were identified. The best genotypes for gluten content (%) were found to be superior by 19.37% -33.11% to the population mean and 18.56% -32.21% to the mean of gluten content performance of the released varieties; protein content 11.13% - 19.55% to the population mean and 12.14% - 20.64% to the mean protein content performance of the released varieties; thousand kernel weight (g) 20.04%-31.13% to the population mean and 9.28%-19.38% to the mean thousand kernel weight performance of the released varieties; zeleny index (ml) 31.42%-52.01% to the population mean and 6.64%-23.35% to the mean biomass performance of the released varieties; vitreousness percentage 33.01%-43.45% to the population mean and 9.57%-18.18% to the mean vitreousness percentage performance of the released varieties (Table 4.4).

A student t-test was used to compare the mean of the top 5% genotypes to the mean of the population for all grain quality traits (Table 4.5). The t-test showed highly significant ($p \leq 0.0001$) differences between means of the selected subsets of the top 5% best genotypes ($\bar{x}$) and the population parameters (μ) for thousand kernel weight (g), moisture content (%), grain yield (t/ha), gluten content (%), protein content (%), zeleny index (ml), and

vitreousness percentage. The study revealed that the top 5% of genotypes had greater relative advantages in all grain quality traits, with differences between them and the population means performances ranging from 5.08% for moisture content to 48.31% grain yield. The top 5% genotypes gave 48.31% greater grain yield (t/ha) than the population as whole. Additionally, the top 5% of genotypes exhibited advantages of 13.61, 24.50, 24.62, 35.62, 39.48, and 48.31 in protein content (%), gluten content (%), thousand kernel weight (g), vitreousness percentage, zeleny index (ml), and grain yield (t/ha) above the mean performances of the population, respectively, indicating the occurrence of various degrees of tetraploid wheat enhancements through selection.

Table 4. 4 Comparison of mean performance of top 5% tetraploid wheat genotypes with mean of the population and released varieties for grain quality traits of tetraploid wheat grown at Sinana and Debrezeit during 2020

Genotypes	Mean	Relative advantage (%) over		Genotypes	Mean	Relative advantage (%) over	
		MP	MRV			MP	MRV
Gluten content				Protein content			
238891	35.26	33.11%	32.21%	6955	15.9	19.55%	20.64%
6102	35.22	32.96%	32.06%	5728	15.44	16.09%	17.15%
7313	34.61	30.65%	29.77%	5504	15.29	14.96%	16.01%
5184	32.84	23.97%	23.14%	Cocorit/71	15.2	14.29%	15.33%
5044	32.37	22.2%	21.37%	5465	15.05	13.16%	14.19%
7020	32.32	22.01%	21.19%	7880	14.95	12.41%	13.43%
DZDW17007	31.99	20.76%	19.95%	ETDW/15DZ#50,ETDW/15DZ#52,……	14.86	11.73%	12.75%
5465	31.91	20.46%	19.65%	214370	14.83	11.50%	12.52%
5504	31.65	19.48%	18.67%	7999	14.83	11.50%	12.52%
ETDW/15DZ#39	31.62	19.37%	18.56%	7313	14.78	11.13%	12.14%
MP	26.49	0	-0.67	MP	13.3	0	0.91
MRV	26.67	0.68	0	MRV	13.18	-0.90	0
Zeleny index				Vitreousness percentage			
5184	63.19	52.01%	23.35%	214370	98.18	43.45%	18.18%
214370	62.89	51.29%	22.76%	7020	94.56	38.16%	13.82%
Ejersa	58.87	41.62%	14.91%	Asasa	93.35	36.4%	12.36%
238891	58.42	40.53%	14.04%	Ejersa	93.04	35.94%	11.99%
DZDW17002	57.61	38.59%	12.45%	7060	92.03	34.45%	10.77%
Tesfaye	57.35	37.96%	11.95%	Werer	91.75	34.06%	10.44%
Werer	56.53	35.99%	10.35%	DZDW17007	91.66	33.93%	10.33%
Yerer	55.64	33.85%	8.61%	DZDW17002	91.32	33.43%	9.92%
DZDW17007	54.64	31.44%	6.66%	5184	91.27	33.36%	9.86%
Selam	54.63	31.42%	6.64%	5638	91.03	33.01%	9.57%
MP	41.57	0	-18.86	MP	68.44	0	-17.62
MRV	51.23	23.24	0	MRV	83.08	21.39	0

Table 4. 5 Comparison of mean performances of selected top 5% genotypes with the population mean performances for grain quality traits of tetraploid wheat germplasm grown at Sinana and Debrezeit during 2020

Traits	Mean of selected genotypes	Population parameter (μ)	Change through selection ($\bar{X}$-μ)	Change as % of a population parameter (μ)	t-test
TKW (g)	42.67	34.24	8.43	24.62	19.41***
GY (t/ha)	4.82	3.25	1.57	48.31	11.68***
GC	32.98	26.49	6.49	24.50	13.95***
MC	11.16	10.62	0.54	5.08	7.39***
Protein	15.11	13.3	1.81	13.61	16.16***
ZI	57.98	41.57	16.41	39.48	17.12***
VITP	92.82	68.44	24.38	35.62	35.25***

The tabulated t-value for 9 degrees of freedom is 4.781 for two-tailed tests and 4.297 for one-tailed tests at 0.1% probability. T = Student t-test; *** = significantly different at 0.1% probability level. TKW= thousand kernel weight, GY= grain yield, GC= Gluten content, MC= moisture content, ZI= zeleny index, and VITP= vitreousness percentage.

4.3.4. Phenotypic and genotypic coefficient of variation: Moisture and protein content showed the lowest PCV and GCV values. The low GCV value for protein content implies that there is minimum variability in protein content (Table 4.6). The low values of PCV, GCV, and ECV for protein content implies that the improvement of protein content by selection is not possible as there is inadequate heritable variation among the genotypes in spite of the low influence of environment (Joshi *et al*., 2018). Intermediate PCV and GCV values were observed for thousand kernel weight (15.1, 11.58) and gluten content (15.30, 10.81). High PCV and intermediate GCV values were observed in zeleny index (23.37, 19.04) and vitreousness percentage (22.82, 13.22). High PCV and GCV values were recorded for grain yield (37.18, 24.03) indicating presence high variability in grain yield and high potential for effective selection (Burton 1953). Similar to the present study, Ganno *et al*. (2017) reported low PCV (6.88) and GCV (4.71) for grain protein content in bread wheat. Similarly, low PCV (4.66) and GCV (3.8) of protein content but as opposed to the present study, medium PCV (16.92) and GCV (10.2) for grain yield and low PCV (8.10) and GCV (6.89) for thousand kernel weight was reported was reported by Taneva *et al*. (2019) in durum wheat genotypes. Kumar (2020) also reported similarly medium and low values PCV and GCV for thousand kernel weight respectively but disagrees on grain yield.

Differences between the value of GCV and PCV respectively considered high, moderate, and low if more than 20%, 10–20%, and less than 10% (Deshmukh *et al*., 1986). The

difference between PCV and GCV was 3.52, 13.15, 2.6, 4.49, 4.33 and 9.6 for thousand kernel weight, grain yield, protein content, gluten content, zeleny index, and vitreousness percentage. Moderate difference between PCV and GCV was observed for grain yield and the difference between PCV and GCV was low for other traits implying that grain yield was substantially influenced by the growing environments while other traits were mainly due to genetic factors and the environment does not play much role in the expression of these traits (Studnicki *et al*., 2016).

Environmental coefficient of variation was high for grain yield (31.38) implying that the variation among genotypes is not heritable and so, selection is not effective in the improvement of grain yield (Joshi *et al*., 2018); intermediate for vitreousness percentage (19.92), zeleny index (18.42), thousand kernel weight (13.46), and gluten content (12.20) and low for protein content (6.73). Genotype by environment interaction coefficient of variation was high for grain yield (33.42) implying the performance of grain yield varies across locations and that there is effect of genotypes, environment, and interaction of genotype and environment in the expression of this trait and low for thousand kernel weight (9.85), vitreousness percentage (9.96), zeleny index (9.21), gluten content (6.10), and protein content (3.36) showing these traits are stable across locations and that the influence of the environment is minimal on these traits. .

4.3.5. Estimate of broad sense heritability (H^2) and genetic advance

H^2 values <40%, 40-80%, and > 80% were categorized as low, medium, and high, respectively (Mesele *et al*., 2015). Estimates of heritability (H^2) ranged from 33.57 % for vitreousness percentage to 66.36 % for zeleny index (Table 4.6). Grain yield (41.78), vitreousness percentage (33.57), and gluten content (49.91) exhibited less than 50% heritability while others remaining traits showed greater than 50% heritability. Medium heritability was recorded for zeleny index (66.36%), moisture content (62.5%), thousand kernel weight (58.85), grain yield (41.78), and gluten content (49.91) and vitreousness percentage (33.57) showed low heritability. Medium broad sense heritability was also reported by Taneva *et al*. (2019) for thousand kernel weight (72) and protein content (67) while low value was reported for grain yield (36).

Table 4. 6 Variability, heritability, and genetic advance of tetraploid wheat germplasm based on grain quality traits

Parameters	TKW	GY	Moisture	Protein	Gluten	ZI	VITP
GV	15.73	0.61	0.05	0.78	8.2	62.64	81.9
EV	21.25	1.04	0.07	0.8	10.45	58.63	185.77
GEIV	11.38	1.18	0.032	1.05	11.24	34.19	231.32
PV	26.73	1.46	0.08	1.51	16.43	94.39	244
GCV	11.58	24.03	2.11	6.64	10.81	19.04	13.22
PCV	15.1	37.18	2.66	9.24	15.3	23.37	22.82
ECV	13.46	31.38	2.44	6.73	12.2	18.42	19.92
GEICV	9.85	33.42	1.22	3.36	6.1	9.21	9.96
Heritability	58.85	41.78	62.5	51.66	49.91	66.36	33.57
GA	18.27	31.94	3.42	9.81	15.7	31.89	15.75
GAM	16.59	26.91	2.85	8.44	13.55	29.34	12.97
The difference between PCV and GCV	3.52	13.15	0.55	2.6	4.49	4.33	9.6

NB: TKW= thousand kernel weight (gm), GY= Grain yield (t/ha). ZI= Zeleny index (ml), EV= Environmental Variance, GV= Genetic Variance, VITP= vitreousness percentage, GCV= Genotypic Coefficient of Variation, .PCV= Phenotypic Coefficient of Variation, ECV= Environmental Coefficient of Variation, GEICV= Genotype by Environment interaction Coefficient of Variation, GAM= Genetic Advance as percentage of Mean, GA= Genetic Advance, PV= Phenotypic Variance, and GEIV= Genotype by Environment interaction variance.

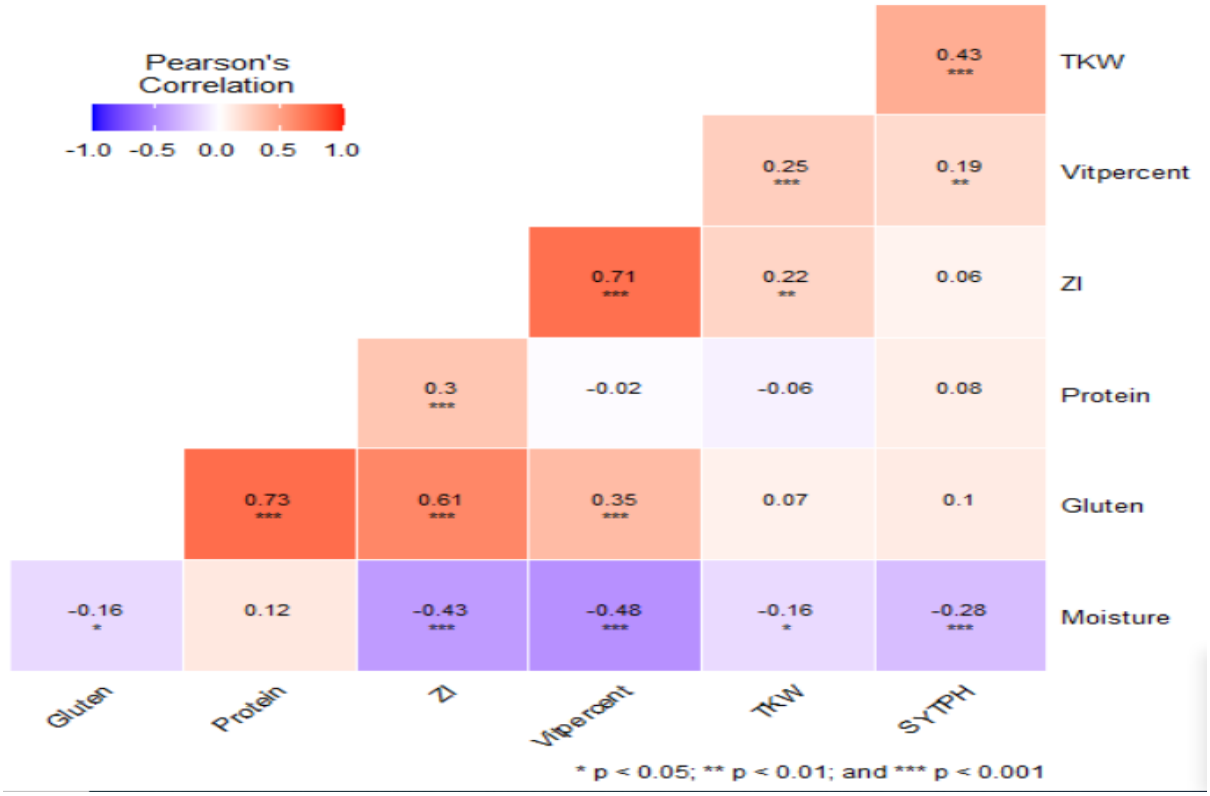
Heritability and genetic advance decide the success of selection in breeding programs. Genetic advance was lowest for moisture content (3.42) and protein content (9.81); intermediate for thousand kernel weight (18.27), gluten content (15.70), and vitreousness percentage (15.75) and high for grain yield (31.94), and zeleny index (31.89). Genetic advance as percentage of mean was low for moisture content (2.85) and protein content (8.44); intermediate for thousand kernel weight (16.59), gluten content (13.55), and vitreousness percentage (12.97) and high for grain yield (26.91) and zeleny index (29.34). In line with the present study, medium value of broad sense heritability (46.86) and low value of GAM (6.65) was reported by Mesele *et al.*(2015) for grain protein content. The present work agrees with Taneva *et al.*(2019) on low and medium GAM values for protein content (6.44) and TKW (12.04) respectively but disagrees on GAM value for grain yield (12.61). Medium heritability and high genetic advance were recorded for zeleny index. Heritability estimates with genetic advance are more useful in predicting the gain under selection than heritability alone (Johnson *et al.*, 1955). Vitreousness percentage had low heritability coupled with medium genetic advance while protein content had low

heritability combined with low genetic advance indicating the lower success in selection for these traits.

4.3.6. Correlation analysis

The result of Pearson correlation coefficient was given in (Table 4.7). Grain yield showed highly significant positive correlation with thousand kernel weight (0.43***) and vitreousness percentage (0.19**). Even though not significant, grain yield also showed positive correlation with seed colour (0.1), protein content (0.08) and gluten content (0.1). There was highly significant negative correlation between grain yield and moisture content (-0.28***). Mesele *et al*. (2015) reported similarly significant positive correlation of seed yield and thousand kernel weight while the report disagrees on grain yield and grain protein content which was significant negative correlation.

Table 4. 7 Pearson's correlation analysis for the eight grain quality traits of 174 landraces along with 22 improved varieties



Remarks: ZI= zeleny index, Vitpercent= vitreousness percentage, TKW= thousand kernel weight, and SYTPH= seed yield in tonnes per hectare.

Thousand kernel weight showed highly significant positive correlation with vitreousness percentage (0.25***) and zeleny index (0.22**). Thousand kernel weight also showed positive correlation with gluten content (0.07) and a negative correlation with protein

content (-0.06) and moisture content (-0.16). In line with this study, Amiri *et al.* (2018) also reported that grain yield and thousand kernel weight had a significant positive correlation.

Vitreousness percentage showed highly negative correlation (-0.48***) with moisture content and a negative, but not significant, correlation with protein content (-0.02). Vitreousness percentage had a highly positive correlation with zeleny index (0.71***) and gluten content (0.35***).

Zeleny index had a highly significant positive correlation with protein content (0.3***) and gluten content (0.61***) and a highly significant negative correlation with moisture content (-0.43***). Protein had a highly significant positive correlation with gluten content (0.73***) and a positive correlation with moisture content (0.12). Gluten showed a negative correlation with moisture content (-0.16). Modoran *et al.*(2010) also reported that there was a significant positive correlation between protein content and zeleny index of wheat.

4.3.7. Principal Component analysis

The first principal component accounted for 34.12% of the total variation while the second component accounted for 20.65 % of the total variation (Table 4.8; Figure 4.2). The third, fourth, and fifth components respectively accounted for 14.35%, 12.18% and 8.85% of the total variation.

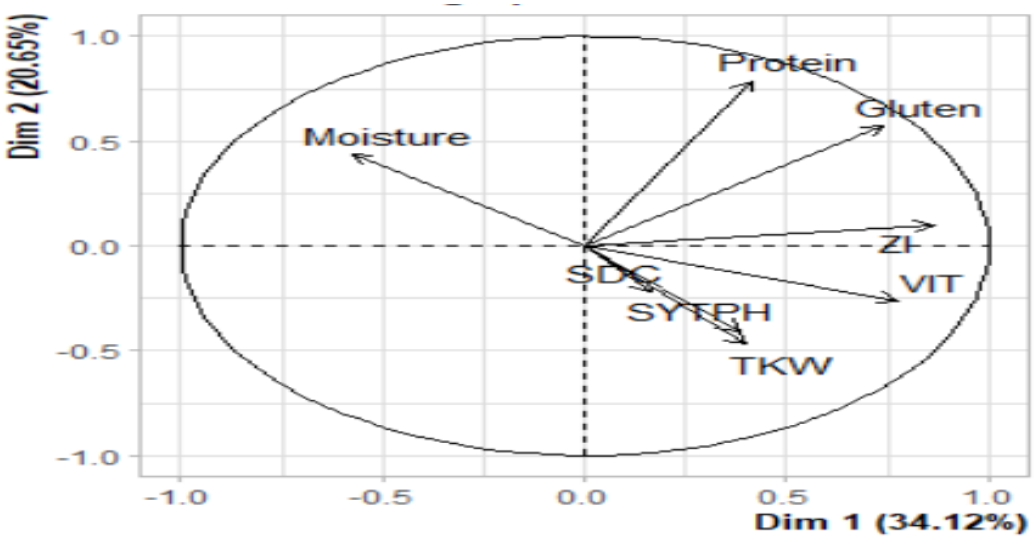


Figure 4. 2 Principal component analysis of eight grain quality traits of174 tetraploid wheat along with 22 improved varieties. Remark: ZI = Zeleny index, VIT = Vitreousness percentage, SDC = seed color, SYTPH = Seed yield (t/ha).

The first three principal components with Eigen values greater than one combined explained 69.11% of the total variation and 90% of the total variation was contributed from the first five components. The variation in the first principal component was mainly contributed by zeleny index (ml), vitreousness percentage, and gluten content. Protein content, gluten content, and thousand kernel weight were the traits which contributed to the variation in the second principal component. Grain yield and thousand kernel weight contributed to the variation in the third principal component.

Table 4. 8 Principal component analysis for the eight grain quality traits of 174 landraces along with 22 improved varieties

Traits	PC1	PC 2	PC 3	PC 4	PC 5	PC 6	PC.7	PC8
TKW	5.751907	13.39591	23.52108	1.54E+00	36.7373	1.72E+01	1.449362	4.23E-01
GY	5.330578	9.920427	39.28342	6.62E-01	19.67339	1.97E+01	4.151279	1.27E+00
SC	0.9953	2.938407	0.148351	9.43E+01	0.060149	4.82E-01	0.182964	9.38E-01
GC	20.03043	19.78378	1.020913	1.04E+00	0.087558	1.29E+00	8.497928	4.83E+01
MC	12.13591	11.67612	7.376306	2.12E+00	32.91308	3.12E+01	1.358868	1.20E+00
PC	6.194237	37.4747	10.58533	4.76E-04	1.889473	8.53E-01	0.553678	4.24E+01
ZI	27.47661	0.623203	7.562789	3.01E-01	5.082611	0.12667	58.82629	8.64E-04
VITP	22.08502	4.187457	10.50181	8.69E-02	3.556432	29.14093	24.97964	5.46E+00
EV	2.729362	1.652106	1.147653	0.974125	0.708115	0.425476	0.192959	0.170205
PV	34.11702	20.65133	14.34566	12.17656	8.85144	5.318449	2.411983	2.127562
CPV	34.11702	54.76835	69.11401	81.29057	90.142	95.46045	97.87244	100

Remarks: PV= Percentage of variance, CPV= Cumulative percentage of variance, GC= gluten content, MC=moisture content, TKW (gm) =Thousand kernel Weight, PC= protein content, SC = seed color, GY (t/ha) = grain yield, EV= Eigen value, and ZI= Zeleny index

Region of collection PCA

PCA of regions of collections indicated that variances of 36.75%, 27.79% and 25.45% were extracted from the first three principal components with Eigen values greater than one, and 89.99 % of the total variance was explained by these components (Table 4.9; Figure 4.3). Zeleny index, gluten content, protein content, grain yield and seed color were the major contributors to the variation in the first principal component. Moisture and protein content majorly contributed to the variation in the second principal component while thousand kernel weight and grain yield were the major contributors to the variation in the third principal component.

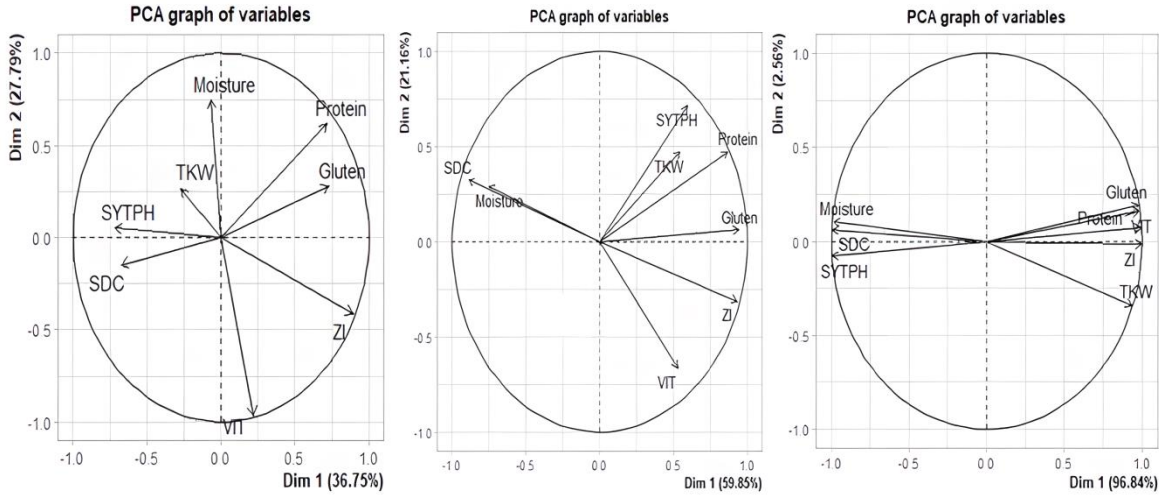


Figure 4. 3 PCA of grain quality traits of tetraploid wheat based on regions of collections (left), altitudinal (middle), and species (right)

Altitudinal based PCA

In the case of altitudinal classes, 59.85%, 21.16% and 14.05% of the total variation were contributed by the first, second and third component respectively (Table 4.9; Figure 4.3). The first and the second principal component together contributed 81% the total variation. The first three components contributed to 95.06% of the total variation. Gluten content, zeleny index, seed color, and protein content were the major contributing traits to the first principal component. Grain yield, thousand kernel weight, and protein content were the major contributing traits to the second principal component. Zeleny index and gluten content majorly contributed to the third principal component.

Species based PCA

The species based PCA revealed that the first principal component alone explained 96.84% of the total variation. The first and second principal component combined explained 99.45% of the total variation. In the case of species, most of the traits contributed almost equally to the variation in the first principal components. Thousand kernel weight was the major contributor to the second principal component (Table 4.9; Figure 4.3).

Table 4. 9 Principal component analysis for the eight grain quality traits of tetraploid wheat based on region of collection, altitudinal, and species

	regional			altitudinal			species		
Variables	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
Thousand kernel Weight (g)	2.478377	3.198995	41.66159	6.148324	13.23262	3.88E+01	11.32371	56.82602	12.89284
Grain yield (t/ha)	17.51924	0.111643	23.21584	7.326835	30.18608	4.13E+00	12.76838	3.081362	9.395733
Seed color	15.47037	0.972528	2.675375	16.47712	6.287666	2.38E-04	12.78349	1.696425	12.92882
Gluten	18.00613	3.485926	9.434822	18.55345	0.24571	7.29E+00	12.37323	17.86194	10.0042
Moisture	0.150433	25.12073	20.9649	11.73878	5.046084	2.68E+01	12.54123	5.225063	37.30914
Protein	17.57694	17.0969	1.49164	15.70102	13.24719	1.13E+00	12.51919	12.72595	8.37532
Zeleny index	27.12996	7.84688	0.549232	18.17716	5.731018	9.80E-01	12.85433	0.123723	8.146695
Vitreousness percentage	1.668552	42.1664	0.006597	5.87731	26.02364	2.09E+01	12.83642	2.459511	0.947244
Eigen value	2.940187	2.223538	2.035606	4.787714	1.692998	1.123675	7.747533	0.205178	0.047289
Percentage of variance	36.75234	27.79422	25.44508	59.84642	21.16248	14.04593	96.84416	2.564729	0.591114
Cumulative percentage of variance	36.75234	64.54657	89.99164	59.84642	81.0089	95.05483	96.84416	99.40889	100

4.3.8. Cluster analysis

The 196 Tetraploid wheat genotypes were grouped into six clusters (Figure 4.4; Table 4.10). The number of clusters varied from 23 in the sixth cluster to 53 in the fourth. The inter-cluster distance varied from 2.16 between cluster 2 and 6 to 4.21 between cluster 4 and 6. The intra cluster distance varied from 1.76 in cluster 3 to 2.18 in cluster 1. The maximum (4.21) inter cluster distance was observed between cluster VI and IV followed by the distance between cluster 5 and 6 (3.44) implying that the accessions in clusters IV and VI are more divergent and that the accessions in these clusters can be used as parents in the breeding program. The maximum intra cluster was observed within cluster 1(2.18) followed by cluster 4 (2.12) (Table 4.11) indicating that the accessions in this cluster are more variable. The cluster analysis revealed that the landraces from the same location do not fall in the same cluster and at the same time landraces from different localities may fall in the same cluster. For example, cluster one was composed of landraces from Shewa, Gojam, Wello, and Arsi. Cluster analysis of the tetraploid wheat showed that there is no distinct regional and altitudinal grouping, which could be due to gene flow or the same selection criteria used by the farmers across regions resulting in similar phenotypes. This result agrees with (Hailu *et al*., 2006; Ayana and Bekele, 1999; Yemataw *et al*., 2018).

Table 4. 10 Clustering of 174 tetraploid wheat landraces along with 22 improved cultivars into two groups using eight grain quality traits

Cluster	Tetraploid wheat genotypes	Number of genotypes	Percentage of total population (n=196)
1	5638,7050,7076,7464,235051,243700, Foka,7880, DZLR27, DZLR31,7666,7578, DZLR7,5168,242789,7069,243703, DZLR23,5618,226694, DZLR25,7007, DZLR9, DZLR32,7999,7313, DZLR28,5504,5214,226241,7690, DZLR18	32	0.16
2	DZLR1,5052, DZLR4,243706,5491,5729,5149, DZLR11,7798,7242,5182, DZLR5,5707,7191,6971, Tesfaye,6975, DZLR19,5666,5140,5492,7078,5600,7647, Ginchi,7195	26	0.13
3	Fetan, Malafia,7507,242781, DZLR3,5534,222427,5576,5584, DZLR16,7014,7209, Quamy, DZLR14, DZLR20, Ejersa,7205,7477, DZLR24,6102,7571,7823,242792,7825, Werer,7762, DZLR36, Asasa,7826,5174,5554,5609,7683, DZLR37, DZLR39,7056, Bekelcha,	37	0.19
4	DZLR2,7020,6968,7568,7581,5526,5197,7532,ld-357, 7046,Arendato,DZLR6, ,DZLR8,5169,5760,DZLR15, 5158,5581,5465,214370,7629,243717,DZLR10,242784,Cocorit/71,6955,5434,242786,242793,5166,5515,DZLR29,7004,Alemtena,7009,243701,5510,5183,242782,DZLR35,DZLR21,5586,7015,5591,5181,5893,5142,7626,242783, DZLR13,7827,6982	53	0.27
5	5373, Selam, Boahi,7378,6987,5913, Mukiye,7295,7713, DZLR12,5369, DZ2018,7075,5728, Gerardo, DZLR26,7832,7150,7060,7673, DZLR33, DZLR30,5171,5342,7664	25	0.13
6	7641,6914,239711,6933,5044,238891, Tob66, DZLR17,5470, DZLR22,5441, Mangudo,7003, Yerer,5172, DZLR34,5057,7063,5635, DZLR38, DZLR40,5216,242790,	23	0.12

NB: numbers or Codes refer to the genotypes used in the study indicated in the lists of the materials used in the experiment.

Table 4. 11 Intra- and inter-cluster distances of tetraploid wheat germplasm based on grain quality traits using average group linkage produced from centroid

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Cluster 1	2.18					
Cluster 2	2.40	1.90				
Cluster 3	2.73	3.24	1.76			
Cluster 4	2.55	3.41	2.75	2.12		
Cluster 5	2.84	2.72	2.44	2.75	1.94	
Cluster 6	3.27	2.16	2.67	4.21	3.44	1.89

Based on region of collections, the landraces from Shewa were clustered with released varieties. The landraces from Wello and Arsi were clustered separately while the landraces from Gojam remained solitary (Figure 4.5). The possible reason for the clustering of released varieties with the landraces from Shewa could be the recollection of improved varieties disseminated to the farmers and under cultivation for longer periods and considered as landraces by farmers (Mengistu *et al*., 2015). The clustering of regions of collections followed a proximity-based pattern (Figure 4.5), suggesting the exchange of genetic material among nearby farmers.

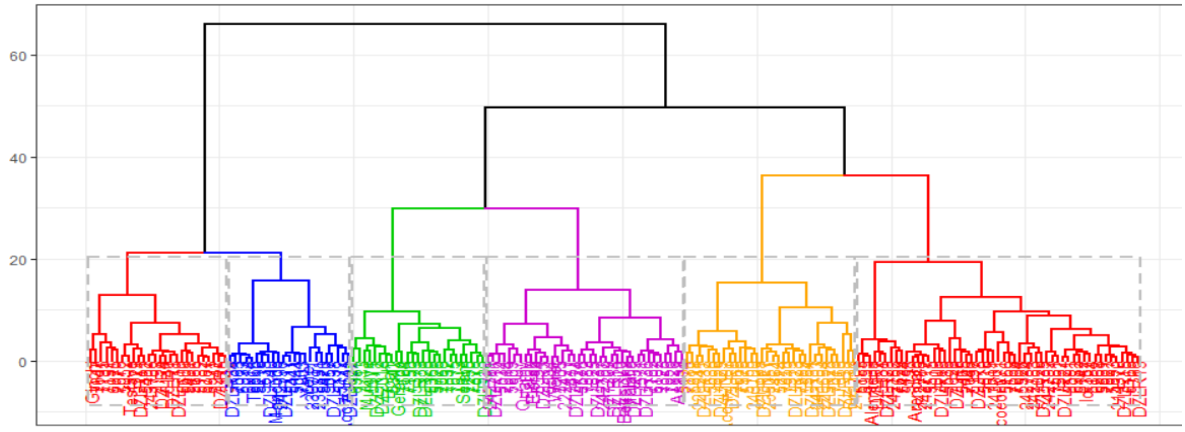


Figure 4. 4 Dendrogram showing the clustering pattern of 174 tetraploid wheat landraces along with 22 improved cultivars based on the eight grain quality traits

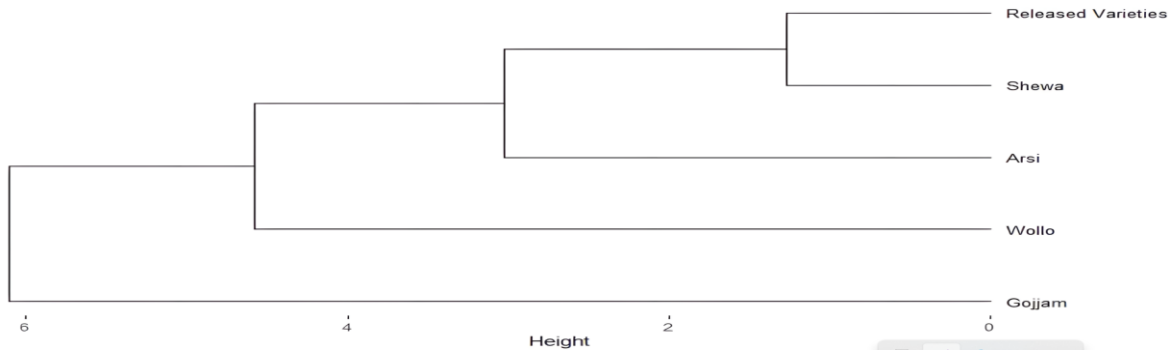


Figure 4. 5 Dendrogram showing the clustering pattern of the regions of collection of Ethiopian tetraploid wheat germplasm based on grain quality traits

The minimum Euclidean distance (1.27) was observed between released varieties and Shewa followed by the distance (2.55) between the landraces from Arsi and released varieties. The maximum Euclidean distance (6.57) was recorded between the accessions from Wello and Gojam followed by the distance (4.78) between the landraces from Wello and Arsi (Table 3.12) implying that the accessions from these regions are more divergent in their grain quality traits.

Table 4. 12 Euclidean Distance of regions of collections of 174 tetraploid wheat landraces along with 22 improved cultivars based on grain quality traits

Region	Shewa	Wello	Arsi	Gojam	Released Varieties
Shewa	0				
Wello	3.19	0			
Arsi	2.63	4.78	0		
Gojam	4.44	6.57	4.23	0	
Released Varieties	1.27	3.33	2.55	4.46	0

Altitudinal classes clustering grouped the landraces into three groups (Figure 4.6). The landraces from altitudinal class 1 (≤1840) and 3 (2021-2200), from altitudinal class 4 (2201-2380), 6(2561-2740) and 8(≥2921) and from altitudinal class 5(2381-2560) and 7(2741-2920) were grouped together. The landraces from altitudinal class 2 (1841-2020) remained solitary. The maximum Euclidean distance 7.291 was recorded between altitudinal class 2 and 3 followed by 7.177 between altitudinal class 1 and 2. The minimum Euclidean distance (1.489) was observed between altitudinal class 4 and 6 (Table 4.13). Clustering of the accessions based on the means of the altitudinal classes indicates that the performance of wheat is influenced by the environment. This result agrees with the result of Hailu *et al*. (2006) on tetraploid wheat from Ethiopia. Species clustering grouped

Triticum aethiopicum and *Triticum durum* grouped together while *Triticum turgidum* was in a separate group. *Triticum polonicum* remained solitary (Figure 4.7). The clustering pattern of species showed that the species are closely related. This result agrees with the result of Hailu *et al*. (2006) on tetraploid wheat of Ethiopia. The maximum and minimum Euclidean distance (5.78, 0.686) was observed between *Triticum turgidum and Triticum polonicum* and *Triticum durum* and *Triticum aethiopicum*, respectively (Table 4.14).

Table 4. 13 Euclidean Distance of altitudinal classes of 174 tetraploid wheat landraces along with 22 improved cultivars based on grain quality traits.

Altitude class	class1	class2	class3	class4	class5	class6	class7	class8
class1	0							
class2	7.177	0						
class3	3.685	7.291	0					
class4	2.853	4.977	3.275	0				
class5	4.599	6.514	4.358	3.345	0			
class6	3.697	4.872	2.687	1.489	3.162	0		
class7	3.295	5.951	3.415	2.287	2.141	2.192	0	
class8	3.059	4.839	3.389	1.757	3.047	1.65	1.899	0

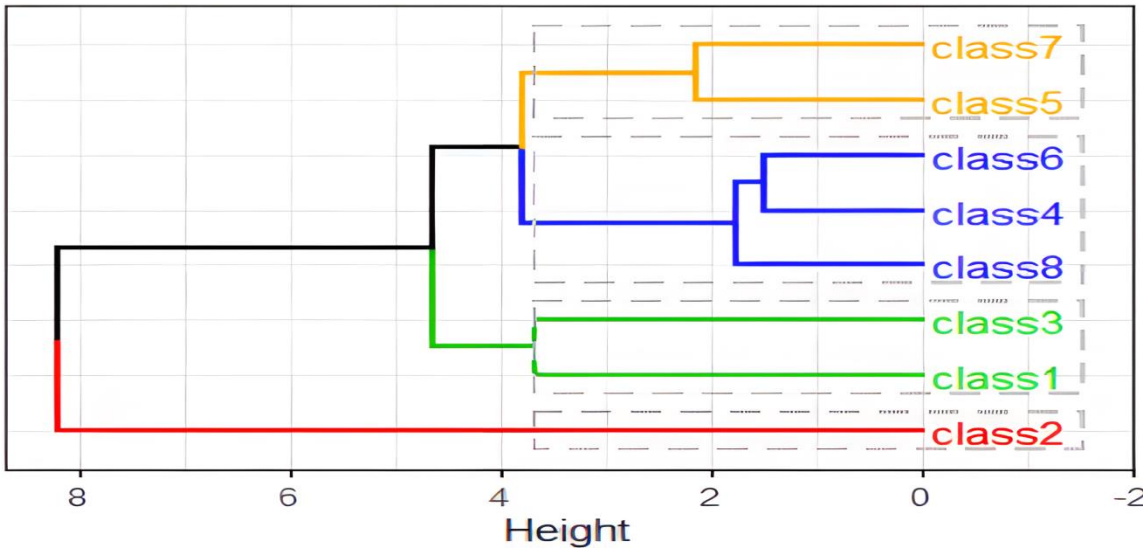


Figure 4. 6 Dendrogram showing the clustering pattern of the altitudinal classes of Ethiopian tetraploid wheat germplasm based on grain quality traits

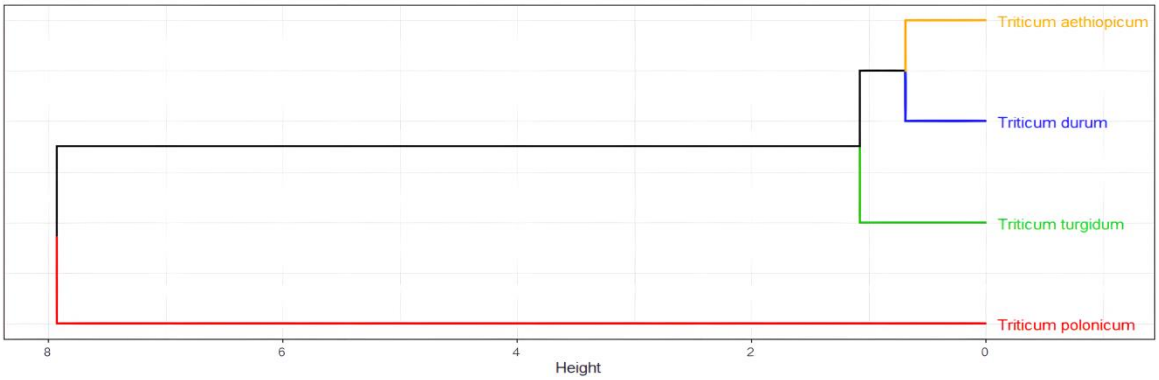


Figure 4. 7 Dendrogram showing the clustering pattern of the species of Ethiopian tetraploid wheat germplasm based on grain quality traits

Table 4. 14 Euclidean Distance of species of 174 tetraploid wheat landraces along with 22 improved cultivars based on grain quality traits

Species	Triticum turgidum	Triticum durum	Triticum aethiopicum	Triticum polonicum
Triticum turgidum	0			
Triticum durum	0.802	0		
Triticum aethiopicum	1.154	0.686	0	
Triticum polonicum	5.78	5.511	5.456	0

4.5. Conclusion

Generally, tetraploid wheat of Ethiopia had huge genetic diversity which could be a potential source for different breeding objectives. Eleven improved varieties (Tob66, Foka, Yerer, Quamy, Mangudo, Ginchi, Asasa, Boahi, Bekelcha, Selam, and Werer) and eleven landraces (7060, 6914, ETDW/15DZ#044, 5510, 7209, 7666, 226694, 5893, DZW17002, 6975, and DZW17007) were selected for production based on their grain quality traits performance. In the present study, the tetraploid wheat accessions are clustered across six groups, which is an indication of high variability. The grouping of accessions by multivariate methods would be of practical value to wheat breeders in that representative accessions may be chosen from different clusters for breeding purposes and to define strategies for further collection of the germplasm. The present study indicates the observed genetic variation is governed by environmental factors. Thus, future germplasm collections should aim to cover collections from as many geographically and climatically different areas as possible, rather than making collections within the region, to enable the country to conserve the huge genetic diversity.

CHAPTER V

5. Genetic diversity in Tetraploid Wheat (*Triticum turgidum spp.*) Germplasm of Ethiopia based on DArTSeq markers

Abstract

Genetic diversity and population structure analysis is vital for germplam conservation and efficient utilization. A set of 188 tetraploid wheat (Triticum turgidum spp) genotypes, including 166 landraces and 22 improved varieties, were genotyped using high density genotyping by sequencing (GBS) by Diversity Arrays Technology sequencing. We investigated the genetic diversity and population structure for the study panel as well as the landraces, separately based on 10349 DArTSeq markers distributed across A and B genomes. The whole tetraploid wheat study panel was clustered into two populations: pop1 and pop2 comprising 150 landraces and three released varieties and 19 released varieties and 12 landraces, respectively. Further population structure analysis of the tetraploid wheat landraces revealed two subgroups of the landraces. Analysis of molecular variance (AMOVA) revealed significant (P<0.001) variation between populations and within populations for both the study panel and the landraces. Total variation within population (88%, 97%) was higher than total variation between populations (12%, 3%) for both study panel and the landraces. The AMOVA result was also verified by the population structure analysis based on genetic differentiation (Fst), which was reported to be little (Fst = 0.028) and moderate (Fst = 0.117) for the landraces and the study panel, respectively implying no significant differentiation among populations. The lower genetic differentiation led to more gene flow in the study panel (Nm=1.891) and in the landraces (Nm=8,612) and this is an indication of more gene exchange preventing genetic differentiation among populations implying a low level of diversity among subpopulations, which is verified by the AMOVA result. The genetic indices of the study panel revealed that pop2 (I=0.523, Ne=1.612, He=0.353, uHe=0.359) was more diverse than pop1 (I=0.415, Ne=1.423, He=0.263, uHe=0.264), which comprised of 98.04% landraces along with 1.96% released varieties; likewise, pop2 (I=0.432, Ne=1.442, He=0.276, uHe=0.277) of the landraces was more diverse than Pop1 (I=1.380, Ne=1.380, He=0.234, uHe=0.236). Furthermore, principal component analysis (PCA) of the tetraploid wheat study panel grouped the released varieties together and the landraces separately; however, PCA of the landraces did not show clear groupings. The first two components together explained 77 % and 68 % of the total variation in the tetraploid wheat study panel and landraces, respectively. The total number of different alleles in pop1 (4762, 4624) and pop2 (4708, 4760) were revealed in the study panel and the landraces, respectively. 127 genotypes were reported to have one or more private alleles at 755 loci, an indication of key adaptive genes at these loci to be exploited in breeding program. The landraces (ETW115 and ETW135) and the released varieties Tesfaye) and Bekelcha could be used in a crossing program owing to their maximum genetic distance. In conclusion, high genetic diversity was detected in Ethiopian tetraploid wheat germplasm, which could be utilized for future wheat breeding program. The higher genetic variation observed within populations than between populations reveal that more attention should be given to individual accessions within populations to explore the existing genetic diversity.

Key words: DArTSeq markers, Genetic diversity, Population structure, and Tetraploid wheat

5.1. Introduction

Wheat is an important market-oriented commodity and a major source of income for many wheat growers in Ethiopia, which is the largest wheat producer in sub-Saharan Africa and a center of diversity for the cultivated tetraploid wheat (Haile *et al*., 2013; Vavilov 1929) that is rich in genetic resources. Despite the existing genetic resources, the average national yield of wheat is low (Haile *et al*., 2013) due to different yield limiting biotic and abiotic factors. Thus, Ethiopia is importing wheat due to the huge gap between production and consumption (Brasesco *et al*., 2019; Abate 2018).

Tetraploid wheat is genetically and morphologically diverse and their evolution under domestication has not been fully elucidated (Oliveira *et al*. 2012). Therefore, assessing the genetic diversity of tetraploid wheat landraces is crucial to address the different breeding purposes such as stable and high grain yield, grain quality, and rust resistance. Knowledge of genetic diversity is very important to identify, conserve and efficiently utilize useful materials from the existing landraces (Oliveira *et al*., 2012; Asmamaw *et al.*, 2020). We can get a better knowledge of the collected material using morphological, biochemical, and molecular markers (Teklu and Hammer 2008), which can help us in the assessment of genetic variability among genotypes within landraces, populations, and species (Eticha *et al*. 2005). Information on available genetic resources, their geographical locations, and understanding of their relationships can be used to gain insight into population divergence (Mengistu *et al*., 2016).

Molecular markers are not influenced by environmental conditions and therefore reflect genetic variation more faithfully (Salvador-Figueroa *et al*., 2015). They can be used in many areas of crop breeding programs (Tian *et al*., 2015) as a research tool to analyse existing populations (Bhalla *et al.*, 2017), for assessment of genetic diversity, germplasm identification, genetic linkage, map construction, map-based cloning, marker assisted breeding, gene mapping and for the mapping of complex traits (Nakaya and Isobe 2012; Meuwissen *et al.*, 2001). Many attempts were made by many authors to study the molecular diversity of Tetraploid wheat of Ethiopia using microsatellites (Teklu *et al*., 2006); Alamerew *et al*., 2004; Geleta and Grausgruber 2012) and microsatellites and markers linked with stem rust resistance (Haile *et al*., 2013), ISSR (Hailu et al. 2005), SNP markers (Negisho *et al*., 2021; Alemu *et al*., 2020), and SSR (Asmamaw *et al.*, 2020). Zang *et al*. (2017) did characterization and identification of sr13 gene that confer

resistance to stem rust Ug99 in tetraploid wheat. Mengistu *et al*. (2016) and Liu *et al*. (2017) did genome wide association study on some durum wheat landraces. Molecular diversity studies indicated that there is great potential in the Ethiopian tetraploid wheat germplasms. However, Negisho *et al*. (2021) indicated that only small portion of the collected landraces are studied. Still the potential of the existing landraces collected and maintained by Ethiopian Biodiversity Institute was not fully studied to effectively utilize the existing germplasm for different economically important traits. 174 landraces of tetraploid wheat are purified and maintained at Sinana Agricultural Research Center and Debrezeit Agricultural Research Center. Therefore, it is crucial to assess the genetic diversity and population structure of these landraces using molecular markers.

Numerous DNA-based genetic markers were developed and contributed greatly to our current understanding of genome organization and genetic variation, but they are constrained by their dependence on gel electrophoresis. To overcome such restrictions, hybridization-based methods using nucleic acids immobilised on solid-state surfaces have been developed. DNA chips have been developed to analyse genotypes for single nucleotide polymorphisms (SNP) but the technology depends on intensive genomic sequencing and a high cost of analysis that cannot be matched in agriculture or basic research. A technology with similar power but with a much greater breadth and much lower cost for DNA polymorphism analysis called Diversity Array Technology (DArT) was reported. Genetic marker analysis through Diversity Arrays Technology offers a low-cost high-throughput, robust system with minimal DNA sample requirement capable of providing comprehensive genome coverage even in organism without any DNA sequence information (Jaccoud *et al*., 2001). Diversity Arrays Technology (DArT) can detect any type of DNA variation at several hundred genomic loci in parallel without relying on sequence information (Wenzl *et al*., 2016) and has been applied successfully for polyploidy and complex genomes such as tetraploid and hexaploid wheat (Robbana *et al*., 2019). Therefore, this study was initiated to assess genetic diversity and population structure of tetraploid wheat landraces of Ethiopia based on DArTSeq markers.

5.2. Materials and Methods

5.2.1. Planting materials

The study panel comprised 188 tetraploid wheat accessions, including 22 released varieties and 166 landraces collected from major wheat-producing areas of the country including Bale, Arsi, Shewa, Jimma, Agaw, Gondar, Gojjam, Tigray, and Wello (Figure 2.1) and detail passport data available in the online resource (Annex 1). The landraces were collected and maintained by Biodiversity Institute of Ethiopia and the utilization of the landraces for research purpose is with the consent of the Institute (IBC 2012).

5.2.2. DNA extraction and genotyping

Leaf samples from two week old seedlings of wheat of each accession grown in greenhouse at Holeta Biotechnology Research Centre was oven dried overnight at 50^{o}c, collected into 96 deep well sample collection plates, and sent to SEQART Africa located in Kenya for high density genotyping by sequencing (GBS) by Diversity Arrays Technology sequencing (DArTseq^{TM} technology), which target active genes, for molecular characterization. DNA extraction was done using Nucleomag Plant DNA extraction kit. The genomic DNA extracted was in the range of 50-100ng/ul. The quality and quantity of genomic DNA were checked on 0.8 agarose. Libraries were constructed according to Kilian *et al*. (2012). DArTSeq complexity reduction through digestion of genomic DNA using a combination of PstI and HpaII enzymes and ligation of barcoded adapters and common adapter followed by PCR amplification of adapter-ligated fragments. Libraries were sequenced using Single Read sequencing runs for 77 bases. Next generation sequencing was carried out using Hiseq2500.

5.2.3. Quality control, SNP calling, and filtering

A sequencing Analysis Viewer was used to check the technical quality of the sequencing. DArTsoft14 software implemented in the KDCompute plungin platform, developed by the Diversity Arrays Technology was used for scoring of DArTSeq markers (http://www.kddart.org/kdcompute.html) and computed several quality parameters for each DArTSeq and SNP markers. The DArTSeq markers generated were aligned to model reference Wheat_ChineseSpring04, which was obtained from the International Wheat Genome Sequencing Consortium (IWGSC) database (https://urgi.versailles.inra.fr/download/iwgsc/) at a minimum base identity of 90% and e-

value of 5e-10. The markers were scored in a binary fashion (1/0), indicating the presence or absence of a marker in the genomic representation of each sample as described by Akbari *et al*. (2006). Marker quality was also maintained by removing monomorphic markers and those with lower call rates.

SNPs with low-quality clustering and those with minor allele frequency (MAF) less than 5% was filtered and removed. Furthermore, the dataset was filtered using a 10% cut-off for missing data in the accessions.

5.2.4. Population structure and genetic diversity analysis

A population stratification of the study panel was visualized by principal component analysis (PCA) using the KDCompute plugin system version 1.01 (https://kdcompute.seqart.net/kdcompute/plugins). Population structure was estimated with STRUCTURE 2.3.4 software (Evanno *et al*., 2005). The Bayesian (maximum likelihood) model-based clustering algorithm (Markov Chain Monte Carlo (MCMC) implemented in STRUCTURE software (Falush *et al.,* 2007) was used to investigate the population structure and assign individual accessions to certain subpopulations based on genetic data. The Evanno method (Evanno *et al*., 2005) was used to determine the likely number of subpopulations via the STRUCTURE HARVESTER program (Earl and vonHoldt 2012). The neighbour-Joining tree was generated from Euclidean distance based on Ward.D method using a plugin system version 1.01 (https://kdcompute.seqart.net/ kdcompute/plugins). Furthermore, a kinship (K) matrix was computed using the method of VanRaden (2008).

Genetic diversity was analysed based on Nei's gene diversity, which is an estimate of the probability that two randomly chosen alleles from the population are different and polymorphism information content (PIC), which reflects the probability of polymorphism between two random samples in the germplasm. Genetic parameters (numbers and percent of polymorphic loci, polymorphism information content (PIC), effective number of alleles, Nei's gene diversity, minor allelic frequency (MAF), and observed heterozygosity) were calculated using Power Marker (Liu and Muse 2005) and GenAlex (Peakall and Smouse 2012). Gene diversity (H) and polymorphic information content (PIC) for each locus were calculated from allele frequencies using

$$PIC = 1 - \Sigma P_i^2\ pj^2 - \Sigma 2P_i^2\ Pj^2$$

and $H = 1 - \Sigma Pi^2$

formulas, where p_i and p_j are the frequencies of the i^{th} and j^{th} alleles, respectively.

5.2.5. Analysis of molecular variance (AMOVA) and genetic diversity indices

Genetic distance between populations was determined using Nei's Genetic Distance (Ne 1978) based on the number of populations k. Analysis of molecular variance (AMOVA) allowed hierarchical partitioning of genetic diversity among populations and within populations (Meirmans and Liu 2018). Analysis of molecular variance (AMOVA) was carried out using GenAlex (Peakall and Smouse 2012) to partition the molecular diversity between and within the populations (k), generated by Structure analysis. The genetic differentiation (FST), which is defined as a standardized measure of the genetic variance among populations was calculated to provide a measure of total genetic divergence between populations (Boucher 2011). The FST value (Nei 1977) of each subpopulation provides an estimate of the degree of fixation of alleles within the subpopulation. The He (Ne 1978) describes the average distance between individuals within the same population, where values close to 0 indicate that the individuals within the population are genetically identical and values close to 1 indicate high genetic variability. Gene flow (Nm) among populations was calculated based on FST as:

$$\mathbf{Nm} = \left[\left(\frac{1}{\mathrm{F_{ST}}}\right) - 1\right]/4$$

Moreover, Shannon's Information Index (I), expected heterozygosity (He), unbiased heterozygosity (uHe), and the percentage of polymorphic loci (PPL) were calculated as follows:

$$I = -1 * \Sigma[Pi * Ln(Pi)],$$

He=1-ΣP_i^2 , uHe = $[\frac{N}{N-1}]$*He.

Where Pi is the frequency of its alleles for the population and ΣPi^2 is the sum of squared population allele frequencies and, $PPL = \Sigma Pi/N$, where Pi is the proportion of loci polymorphic in a population and N the number of populations.

The genetic diversity indices such as the number of different alleles (Na), number of effective alleles (Ne), and Shannon's Information Index (I) were estimated using GeneAlEx (Peakall and Smouse 2012).

5.3. Results and Discussion

Ethiopian tetraploid wheat hold extremely high potential of genetic diversity (Haile *et al*. 2013 ; Muleta *et al*., 2017; Zegeye *et al.*, 2020), genetically unique (Wang *et al*., 2007), and are highly prized for their potential value as source of variation in crop improvement programs for different economically important traits (Mengistu *et al*., 2016; Hardon 1992; Liu *et al*., 2017) such as earliness, disease and pest resistance, nutritional quality, resistance to drought and other stress conditions, and characteristics useful in low-input agriculture (Hardon 1992). Hence, understanding the level and structure of the genetic diversity of tetraploid wheat is a prerequisite for the conservation and efficient use of the available germplasm for plant breeding (Laidò *et al*., 2013). Accordingly, the genetic diversity in tetraploid wheat of Ethiopia was conducted using Diversity Array Technology (DArT), which has been applied successfully for polyploidy and complex genomes such as tetraploid and hexaploid wheat (Robbana *et al*., 2019). DArTSeq markers were used in the present study.

5.3.1. SNP markers distributions

17111 DArTSeq markers were generated from the study panel (Fufa *et al*., 2023) out of which 16308 were polymorphic and 11943 (69.80%) SNPs were with known positions. The observed number of SNPs in the present study is comparable with the number of SNPs (18788, 11919) reported by Alemu *et al*. (2020) and Negisho *et al*. (2021), respectively in durum wheat of Ethiopia. Among 17111 DArTSeq markers, 10349 SNPs, including 10220 (98.75%) polymorphic and 129 (1.25%) monomorphic markers, were distributed across A (48.31%) and B (51.69%) genome (Figure 5.1). The highest (995) and the lowest (503) number of polymorphic SNPs were on chromosome 2B and 4B, respectively. Considering both A and B genomes together, the highest number of polymorphic markers were on the second (1799) chromosome followed by the seventh (1718) chromosome. Genome B scored high number of SNPs as compared to the genome A. The second and seventh chromosome scored high number of markers. B genome scored high number of markers in the present study. In line with this study, (Alemu *et al*., 2020; Negisho *et al*. 2021) reported high number of SNPs in B genome as compared to A genome in Ethiopian durum wheat study panel, Zhang *et al*. (2011) reported high number of SNPs in B genome than in B genome followed by D genome in common wheat cultivars of China, and Kumar *et al*. (2019) also reported high number of SNPs in B genome of Indian spring wheat. The lowest

number of polymorphic SNPs (1107) was on the fourth chromosome. However, the highest number of polymorphic SNPs were scored on chromosome 2B (995) followed by chromosome 7A (943). Similar to this study, Kumar *et al*. (2019) reported high number of SNPs on chromosome 2B in India spring wheat. The reason for the observation of high level of polymorphic markers in B genome than A genome in the present study as well as other similar studies on wheat genome might be due to the fact that B genome is the most differentiated and divergent from the ancestral donors among the wheat genomes as evidenced from the cytogenetic and molecular investigations (Blake *et al*., 1999; Maccaferri *et al*., 2003; Zhang *et al*., 2002; Zohary and Feldman 1962) and that all Triticum species share A genome which is identical with the chromosomes of diploid *triticum boeoticum* (Zohary and Feldman 1962).

After filtering for MAF < 0.05 (Figure 5.2) and Heterozygosity = 0.2, 2381 markers (0.35≤ PIC ≤ 0.50 and call rate ≥ 0.75) were retained for molecular diversity analysis given in the online resource (Fufa *et al*., 2023). The accessions were also filtered for missing percentage at a cut-off =0.1 and finally 184 genotypes were retained for further molecular analysis. The polymorphic information content of the SNPs ranges from 0.01 to 0.5 in both A and B genomes with an average of 0.18 and 0.19, respectively. The average PIC ranges from 0.17 to 0.21 in genome A and from 0.18 to 0.22 in B genome. The average MAF values ranges from 0.11 to 0.15 in both A and B genomes. Chromosome 1B and 5B scored highest average PIC (0.22) (Table 5.1). The PIC values of the polymorphic SNPs ranges from 0.01 to 0.5 with an average value of 0.19. Zhang *et al*. (2011) reported 0.03 to 0.5 with a mean value of 0.40 in common wheat cultivars of China. The average PIC ranges from 0.17 to 0.21 in genome A and from 0.18 to 0.22 in B genome. In line with this study, Alemu *et al*. (2020) reported 0.18 to 0.22 in A genome and 0.19 to 0.22 in B genome of Ethiopian durum wheat based on high density SNP markers. Likewise, Kabbaj *et al*. (2017) reported PIC value of 0.13 in *Triticum aethiopicum* type in the global panel of durum wheat diversity. Ren *et al*. (2013) reported mean value of PIC from 0.157 to 0.2072 in A genome and from 0.1747 to 0.2225 in B genome in a Worldwide Germplasm Collection of Durum Wheat based on SNP markers. The mean values of gene diversity, heterozygosity, and MAF respectively ranges from 0.59 to 0.60, 0.02 to 0.03, and 0.11 to 0.15 in both A and B genome (Table 4.1). The average values gene diversity, PIC, and heterozygosity were respectively 0.34, 0.29, and 0.08 at the whole genotype level (Table 5.1)

The mean values of gene diversity, heterozygosity, and MAF respectively ranges from 0.59 to 0.60, 0.02 to 0.03, and 0.11 to 0.15 in both A and B genomes. The mean gene diversity of chr1A, chr4A, chr6A, chr7A, chr1B, chr6B, and chr7B was 0.60 while that of chr2A, chr3A, chr5A, chr2B, chr3B, chr4B, and chr5B was 0.59 (Table 4.1).

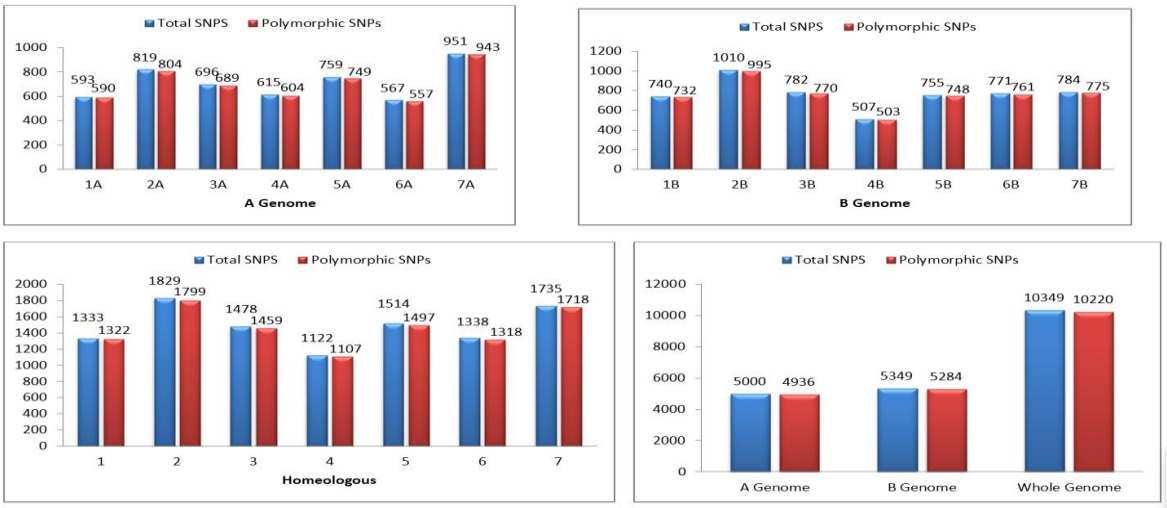


Figure 5. 1 Distribution of 10349 Single nucleotide polymorphisms (SNPs) across the Tetraploid wheat genome. Genome A and B are marked with blue and red colours, respectively

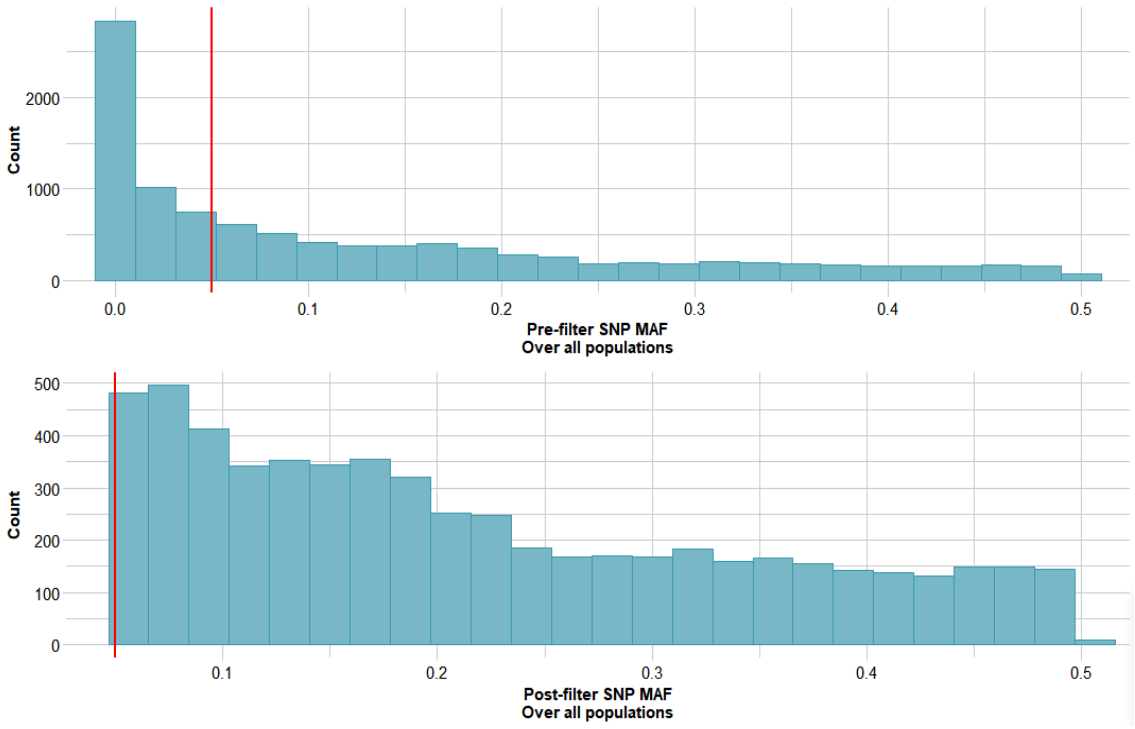


Figure 5. 2 Minor allele frequency distribution of tetraploid wheat study panel before and after filtering

Table 5. 1 The mean values of Minor allele frequency, Gene diversity, heterozygosity, and PIC across chromosome of tetraploid wheat study panel, at species level, regions of collection, and the whole population level

Chromosome	Call Rate	Gene Diversity	Heterozygosity	Average and range of Polymorphic Information Content	Minor Allele Frequency
A Genome					
chr1A	0.72	0.6	0.02	0.21(0.01-0.50)	0.15
chr2A	0.73	0.59	0.02	0.17(0.01-0.49)	0.11
chr3A	0.72	0.59	0.02	0.18(0.01-0.49)	0.12
chr4A	0.72	0.6	0.03	0.17(0.01-0.50)	0.11
chr5A	0.73	0.59	0.03	0.18(0.01-0.50)	0.12
chr6A	0.72	0.6	0.03	0.18(0.01-0.50)	0.12
chr7A	0.71	0.6	0.03	0.21(0.01-0.49)	0.14
B Genome					
chr1B	0.71	0.6	0.03	0.22(0.01-0.49)	0.15
chr2B	0.73	0.59	0.02	0.19(0.01-0.50)	0.13
chr3B	0.72	0.59	0.02	0.18(0.01-0.49)	0.12
chr4B	0.72	0.59	0.02	0.19(0.01-0.50)	0.13
chr5B	0.73	0.59	0.03	0.22(0.01-0.50)	0.15
chr6B	0.71	0.6	0.02	0.18(0.01-0.49)	0.11
chr7B	0.71	0.6	0.03	0.19(0.01-0.49)	0.13

	Major.Allele.Frquency	GenotypeNo	Sample size	Allele No	Gene Diversity	Heterozygosity	PIC
Arsi	0.848401	2.351953	24	2.460731	0.226511	0.075336	0.20048
Gojam	0.81949	2.156237	22	2.252415	0.251714	0.070158	0.214868
Released varieties	0.76448	2.321714	22	2.403192	0.322776	0.091539	0.273177
Shewa	0.78685	2.927761	97	2.950441	0.307587	0.076239	0.266202
Wello	0.805953	2.274675	19	2.364973	0.272785	0.067375	0.232826
Triticum aethiopicum	0.840373	1.98152	14	2.061319	0.223624	0.071459	0.193134
Triticum durum	0.774544	2.856783	58	2.900882	0.323338	0.077865	0.27801
Triticum polonicum	0.960101	1	1	1.079798	0.039899	0.079798	0.029924
Triticum turgidum	0.802953	2.769845	89	2.813104	0.280068	0.072253	0.241865
Mean	**0.77**	**2.98**	**184**	**2.99**	**0.34**	**0.08**	**0.29**

5.3.2. Population structure and genetic relationships

Population structure and kinship (K) matrix analysis were carried out based on DArTSeq markers. Population structure analysis of tetraploid wheat study panel for the whole tetraploid wheat study panel revealed Δk at k=2, that is, two populations hereafter referred to as pop1 and pop2 (Figure 5.3 a-b and Table 5.2), where 83.15 % (153) were assigned to cluster one and 31 (16.85%) were assigned to cluster two. Pop1 comprised of three released varieties (1.96%) (Quamy, Bekelcha, and Arendato) and 150 landraces (98.04%). Pop2 comprised of nineteen released varieties (61.29%) along with twelve landraces (38.71%) (5168, 'MCD11-20, MCD8-2, ETDW/15DZ019', 5591, 5586, 222427, 5728, 243703, MCD4-4, MCD4-32, ETDW/15DZ#044, MCD3-15, and DZDW17002), which may be admixture or incorrect renaming of the released varieties during the collection of the germplasm (Negisho *et al*., 2021). Row two of Table 5.2 indicates the optimum number of populations in the study panel is two; likewise, Neighbor Joining (NJ) revealed two groups of the study panel (Figure 5.4). A low level of population structure was also supported by the VanRaden (Figure 5.5C), which was in accordance with NJtree and PCA analyses. The kinship coefficient between pairs of 184 accessions ranged from 0.00 to 1.488. Figure 4.5B represents the 3D plots of the first three principal components to depict the samples' relationship in space, the analysis also confirmed the presence of kinship in the association panel (Figure 5.5C), suggesting the importance of using a powerful statistical GWAS model that accounts for the population structure and familial relatedness in the association study. Based on kinship analysis, eleven genotypes were selected for crossing in a breeding program due to their low kinship values between their pairs. Average distance (expected heterozygosity) between individuals in the same cluster was 0.2527 and 0.2256 for cluster 1 and cluster 2, respectively. The allele frequency divergence among populations of the study panel was 0.3164 (Table 5.3). In line with this study, Negisho *et al*. (2021) reported nine landraces clustered along with the released varieties and these landraces could be admixtures. Oliveira *et al*. (2012) also reported admixtures in Tetraploid wheat landraces in the Mediterranean basin. Opposite to this study, Alemu *et al*. (2020) reported three sub-populations in Ethiopian durum wheat. The three released varieties clustered with the landraces may be a selection from the landraces.

The maximum (0.387) and the minimum distances (0.011) in the study panel were observed between ETW115 (7683) and ETW135 ('MCD11-20, MCD8-2, ETDW/15DZ019,……') and ETW104 (235051) and ETW105 (7507), respectively. The maximum (0.553) and minimum

(0.0012) distance in the released varieties, based on the DArTSeq markers, were respectively observed between ETW185 (Tesfaye) and ETW 169 (Bekelcha) and ETW185 (Tesfaye) and ETW187 (Alemtena) (Fufa *et al*., 2023). Alamerew *et al*. (2004) reported the average genetic distance of 0.528 in tetraploid wheat of Ethiopia. In line with this study, Zhang *et al*. (2011) reported 0 to 0.76 Jaccard genetic distance in common wheat of China. Opposite to this study, Kumar *et al*. (2019) reported lower genetic distance (0.019) between populations (sp1 and sp2) of Indian spring wheat.

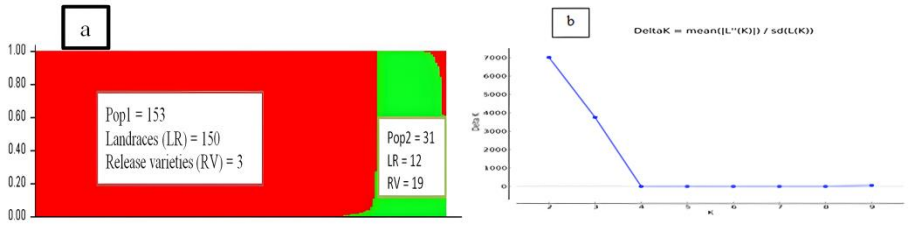


Figure 5. 3 Population structure analysis of tetraploid wheat study panel based on DArTSeq markers (a) Bayesian structure analysis, (b) Structure harvester Evanno's test ΔK at K=2.

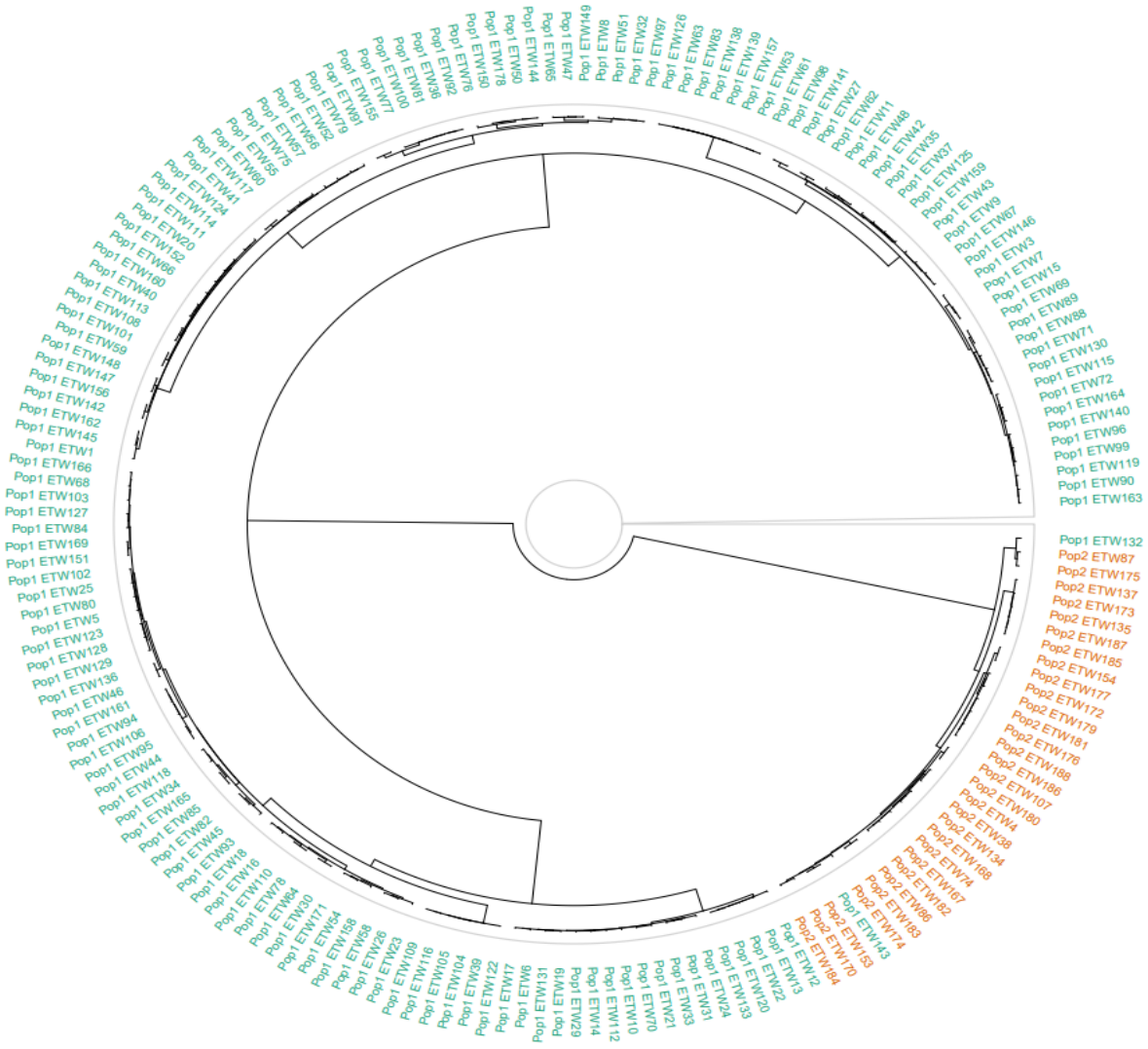


Figure 5. 4 Neighbor Joining (NJ) tree generated from Euclidean distance based on Ward.D method using 2381 DArTSeq markers from 184 tetraploid wheat genotypes. Populations identified in STRUCTURE based on Bayesian structure analysis are shown in green and red for Pop1 and Pop2, respectively

Table 5. 2 Inference of the optimal numbers of clusters existed in Ethiopian tetraploid wheat panel using Delta K Statistics based on DArTSeq markers

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln"(K)|	Delta K
1	5	-398519.3	0.5933	—	—	—
2	5	-298845.3	8.6794	99674	60869	7013
3	5	-260040.3	6.8642	38805	25779	3755.6
4	5	-247014.4	2047.8	13026	2009.8	0.9814
5	5	-235998.4	1432.4	11016	3615.7	2.5243
6	5	-228598	2756.7	7400.4	2925.2	1.0611
7	5	-224122.8	3363.1	4475.2	4212.5	1.2526
8	5	-215435.2	2665.1	8687.6	3684.1	1.3824
9	5	-210431.7	803.34	5003.5	45033	56.058
10	5	-250461.4	93492	-40030	—	—

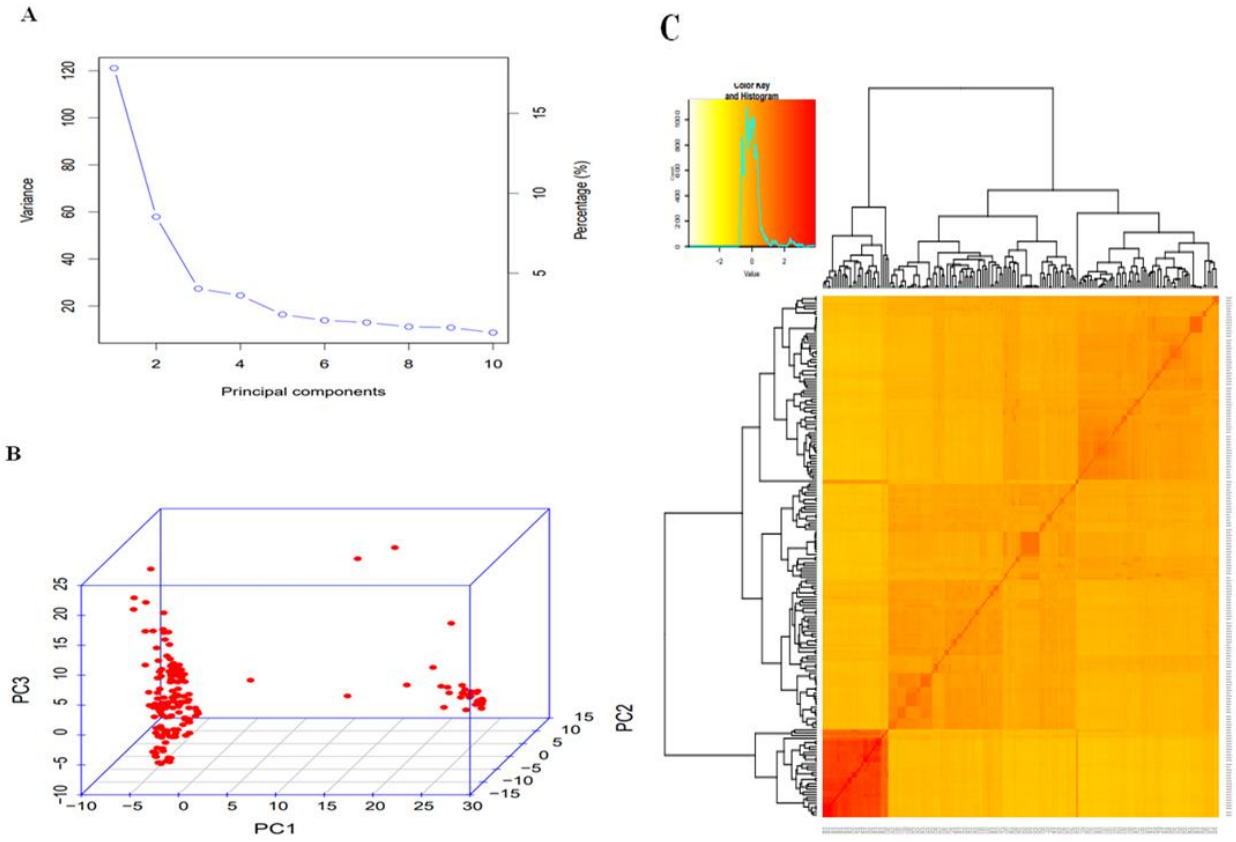


Figure 5. 5 Principal component and genetic relatedness of 184 tetraploid wheat genotypes based on 2,381 genome-wide scanned, high-quality SNPs. **(A)** A screen plot displaying the first 10 principal components with their corresponding fraction of variation explained, **(B)** 3D plots of the first three principal components to depict the samples' relationship in space, and **(C)** Heat map showing the kinship analysis. The kinship values showed a normal distribution (turquoise curve), and orange and red colors represent weak and high kinship relations in the panel, respectively. The resulted clustering tree is indicated outside of the matrix.

Table 5. 3 Population structure values generated using DArTSeq markers for study panel and landraces of tetraploid wheat

	Study panel		Landraces	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2
Overall proportion of membership of the sample in each cluster	83.15	16.85	46.6	53.4
Average distance (expected heterozygosity) between individuals in the same cluster	0.2527	0.2256	0.2809	0.1885
Allele-freq. divergence among pops (Net nucleotide distance), computed	Study panel		Landraces	
using point estimates of P.	Cluster 1	Cluster 2	Cluster 1	Cluster 2
Cluster 1		0.3164		0.1271
Cluster 2	0.3164		0.1271	

Population structure analysis, based on DArTSeq markers, of the tetraploid wheat landraces, on the other hand, revealed Δk at k=2, i.e., two sub-populations (Pop1-1 and pop1-2) of the landraces (Figure 5.6a-b; Table 5.4), which is opposite to the study of (Negisho *et al*. 2021). The yellow highlighted row in Table 5.4 indicates the optimum number of populations in the tetraploid wheat landraces is two. Furthermore, Neighbor Joining (NJ) revealed two groups in the populations of the landraces (Figure 5.7). Sub-population 1-1 and 1-2 comprised seventy-five (Shewa=43, Wello=13, Arsi=10, Gojam=6, and Gonder=2) and eighty-seven (Shewa=52, Arsi = 16, Gojam=11, Wello=6, and Gonder=2) landraces. Average distance (expected heterozygosity) between individuals in the same cluster was 0.2809 and 0.1885, respectively for cluster 1 and cluster 2. The allele frequency divergence among populations of the landraces was 0.1271 between cluster 1 and cluster 2 (Table 5.3). In line with the study of many authors (Alemu *et* al., 2020; Kabbaj *et al*., 2017; Mengistu *et al*., 2016; Negisho *et al*., 2021) our study also revealed higher admixture of accessions between different populations of tetraploid wheat landraces that could arise from the human transfer of genes in history or the existence of gene flow through informal seed exchange in the country or across regions (Negisho *et al*., 2021; Zewdie *et al*., 2004).

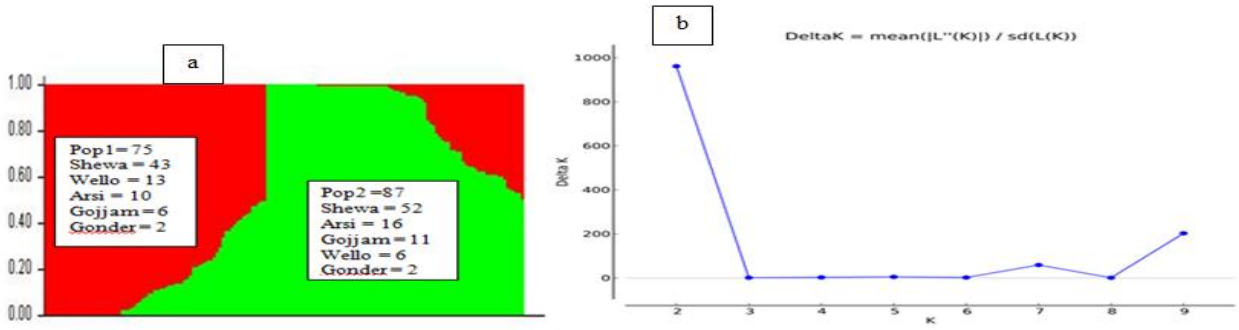


Figure 5. 6 Population structure analysis of tetraploid wheat landraces based on DArTSeq markers. (a) Bayesian structure analysis, (b) Structure harvester Evanno's test ΔK at K = 2. Pop1 and Pop2 represents population one and population two, respectively while sd stands for standard deviation

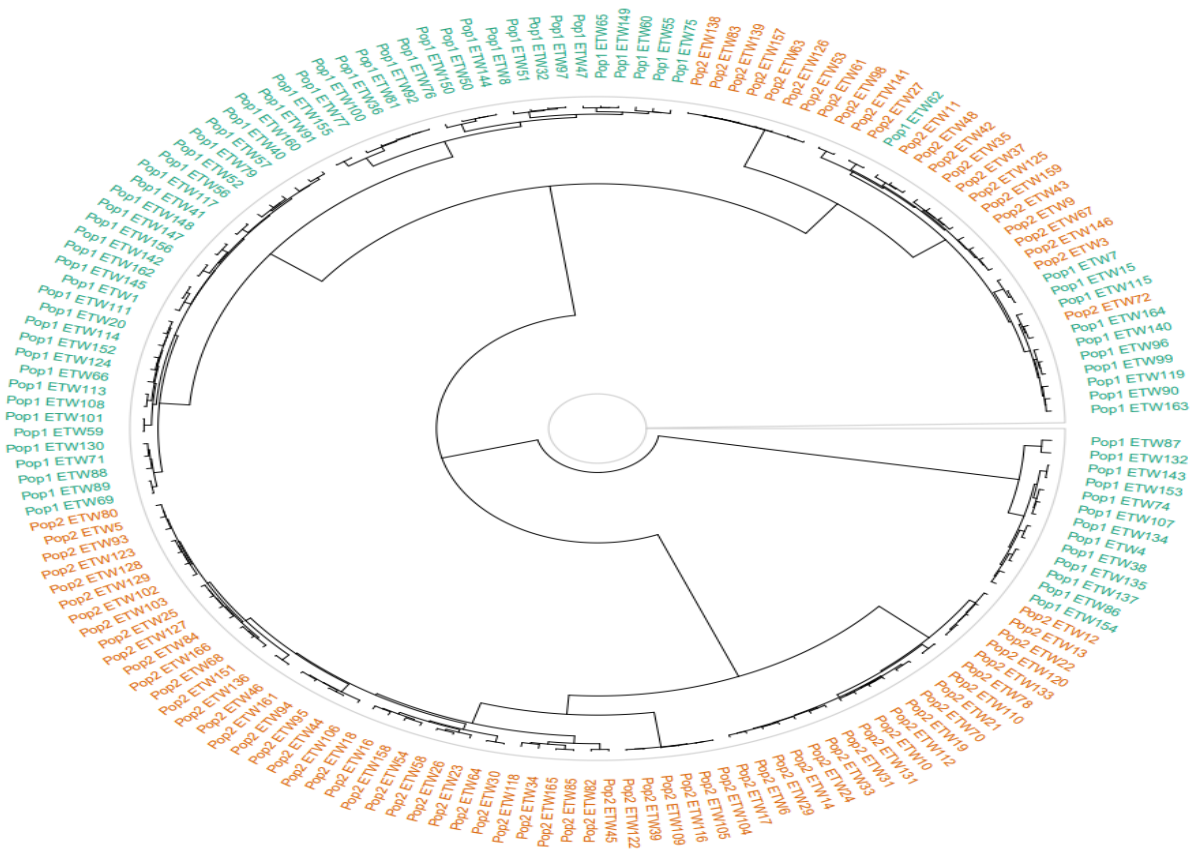


Figure 5. 7 Neighbor Joining (NJ) tree generated from Euclidean distance based on Ward.D method using 2381 DArTSeq markers from 162 tetraploid wheat landraces. Populations identified in STRUCTURE based on Bayesian structure analysis are shown in green and red for Pop1 and Pop2, respectively

Table 5. 4 Inference of the optimal numbers of clusters existed in Ethiopian tetraploid wheat landraces panel using Delta K Statistics based on DArTSeq markers

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	5	-305682.46	4.4903	—	—	—
2	5	-262422.94	10.097	43260	9705	961.18
3	5	-228868.42	12239	33555	15070	1.2313
4	5	-210383.52	2199.1	18485	5823.4	2.6481
5	5	-197722	1627	12662	8158.5	5.0145
6	5	-193219	2511.1	4503	5216.8	2.0775
7	5	-183499.22	1655.7	9719.8	97513	58.895
8	5	-271292.44	204789	-87793	184609	0.9015
9	5	-174476.22	803.63	96816	162919	202.73
10	5	-240579	124313	-66103	—	—

5.3.3. Principal component analysis

Similar to the population structure, principal component analysis (PCA) of tetraploid wheat study panel revealed the separate grouping of the released varieties from the landraces except three released varieties, which were grouped along with the landraces (Figure 5.8); the principal component analysis of the tetraploid wheat landraces showed no clear grouping of the populations of the landraces implying that there is no clear genetic differentiation among the populations of the landraces due to gene flow (Figure 5.9) except thirteen landraces, which are distinctly separated from others. The first two components together explained 77 % (PC1 = 65 %, PC2 = 12 %) and 68 % (PC1 = 48 %, PC2=20 %) of the total variation in the data of tetraploid wheat study panel and landraces, respectively. Likewise, a scree plot, which was used to display the proportion of variation captured by each of the 10 principal components, also showed that the first two principal components (PC1 and PC2) explained the highest proportion of the total variation in the panel (Figure 5.5A). The present study revealed that the genetic diversity within population is higher than the genetic diversity between populations implying that due emphasis should be given to individuals within populations to explore the genetic diversity of tetraploid wheat for conservation and genomic analysis (Negisho *et al*., 2021). In line with this study, Negisho *et al*. (2021) reported higher percentage of variation within populations than between populations of durum wheat study panel. Moreover, Hagenblad *et al*.(2012) observed high within population diversity in their study of the population genetics of gene bank and historical landrace varieties of barley, pea, oat, and rye.

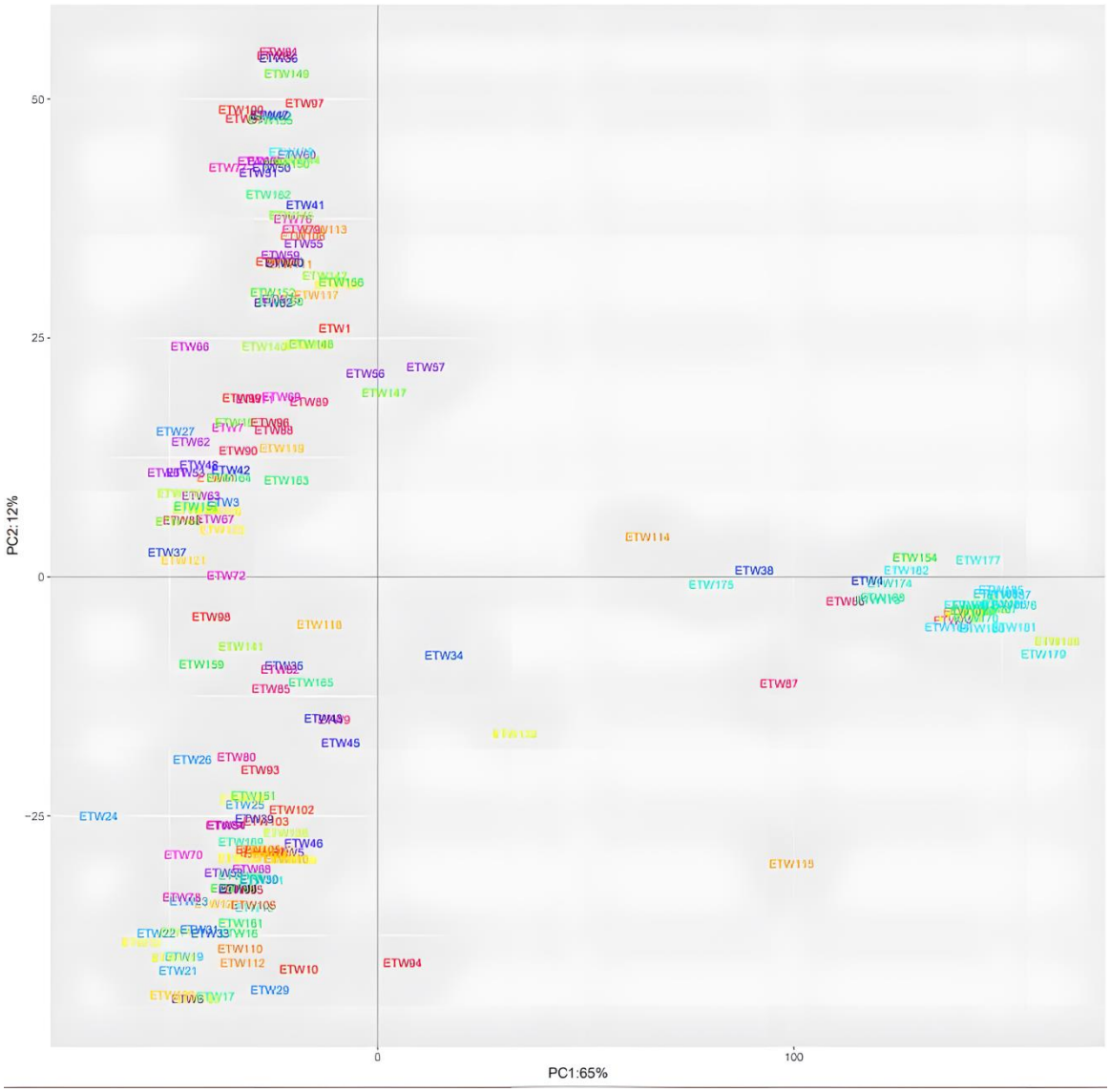


Figure 5. 8 Population structure analysis of 184 tetraploid wheat genotypes based on the principal component analysis clustering as revealed by the first two principal components (PC1 and PC2) produced using 2381 polymorphic DArTseq markers

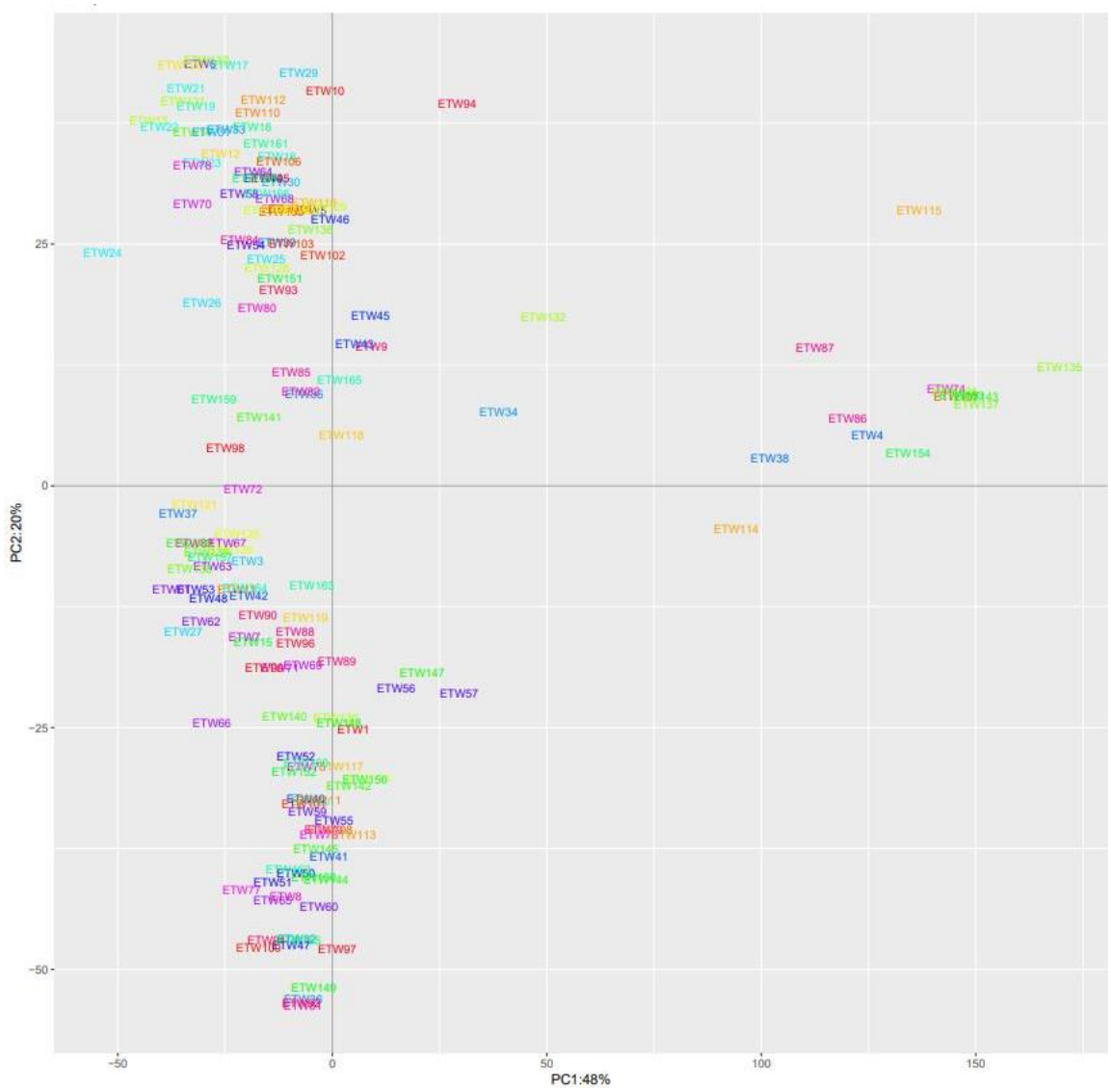


Figure 5. 9 Population structure analysis of 162 tetraploid wheat landraces based on the principal component analysis clustering as revealed by the first two principal components (PC1 and PC2) produced using 2381 polymorphic DArTseq markers

5.3.4. Genetic diversity indices analysis

The genetic indices for the study panel such as I, He, and uHe showed higher values for pop2(I=0.523, Ne=1.612, He=0.353, uHe=0.359, PPL=97.73), which comprised 19 released varieties and 12 landraces, as compared to pop1(I=0.415, Ne=1.423, He=0.263, uHe=0.264, PPL=100), which is comprised of 98.04 % landraces and 1.96% released varieties imply ing that the populations of released varieties (I=0.523, Ne=1.612, He=0.353, uHe=0.359, PPL= 97.73) were more diverse than the landraces populations (I=0.415, Ne=1.423, He=0.263, uHe=0.264, PPL = 100) (Table 5.5). Moreover, the mean gene diversity of released varieties (gene diversity=0.18, PIC=0.15) is higher than that of

the landraces (gene diversity = 0.15, PIC = 0.12). The population analysis of the tetraploid wheat landraces revealed two sub-populations; that is, sub-population 1-1 and sub-population 1-2, which comprised 46.3 % (75) and 53.7 % (87) of the landraces, respectively. Sub-population two showed more diversity (Ne=1.442, I= 0.432, He =0.276, uHe=0.277, and PPL =99.92) than sub-population one (Ne=1.380, I= 0.367, He =0.234, uHe=0.236, and PPL =94.2). Sub-population one was comprised of 43, 13, 10, 6, and 2 landraces from Shewa, Wello, Arsi, Gojam, and Gonder, respectively while sub-population two was comprised of 52, 16, 11, 6, and 2 landraces from Shewa, Arsi, Gojam, Wello, and Gonder, respectively. Opposite to this study, Negisho *et al*. (2021) reported that the populations of the landraces are more diverse than the populations of released varieties. Opposite to the present study, Kumar *et al*. (2019) reported high mean values of genetic indices (I=0.643, He=0.451, and uHe = 0.452 in sp1 and I = 0.648, He = 0.456 and uHe = 0.456 in sp2) in Indian spring wheat. In line with this study, Kumar *et al*. (2019) reported He values of 0.22 (*ssp. carthlicum*) to 0.33 (ssp. Durum) in Indian spring wheat.

Table 5. 5 Mean of different genetic indices of the genetic parameters in populations of tetraploid wheat genotypes

Population of the tetraploid wheat k=2, based on DArTSeq markers								
Population	Sample number	Na	Ne	I	Ho	He	uHe	PPL
Pop1	153	2	1.423	0.415	0.047	0.263	0.264	100
Pop2	31	1.977	1.612	0.523	0.064	0.353	0.359	97.73
	Population of tetraploid wheat landraces k=2, based on DArTSeq markers							
Pop1-1	75	1.942	1.380	0.367	0.041	0.234	0.236	94.20
Pop1-2	87	1.999	1.442	0.432	0.051	0.276	0.277	99.92

Remarks: Shannon's information index (I): I= −1 * ∑[Pi *LnPi], expected heterozygosity (He) or genetic diversity: He = 1 − ∑Pi^2, and Unbiased heterozygosity uHe= [N/(N-1)] * He., Na = No. of Different Alleles, Ne = No. of Effective Alleles = 1 / (Sum pi^2), I = Shannon's Information Index = -1* Sum (pi * Ln (pi)), Ho = Observed Heterozygosity = No. of Hets / N, He = Expected Heterozygosity = 1 - Sum pi^2, uHe = Unbiased Expected Heterozygosity = (2N / (2N-1)) * He, F = Fixation Index = (He - Ho) / He = 1 - (Ho / He), Where pi is the frequency of the ith allele for the population & Sum pi^2 is the sum of the squared population allele frequencies, and PPL = percentage of polymorphic loci.

5.3.5. Allelic patterns across the populations of tetraploid wheat

Allelic patterns across the populations of tetraploid wheat based on DArTSeq markers for the tetraploid wheat study panel and the landraces was given in Figure 5.10 and Table 5.6.

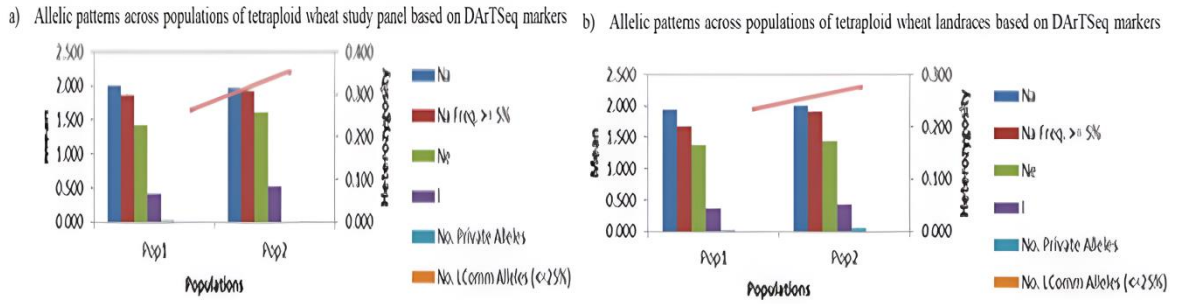


Figure 5. 10 Allelic patterns across the populations of tetraploid wheat study panel and landraces based on DArTSeq markers. Remarks: Na = No. of Different Alleles, Na (Freq >= 5%) = No. of Different Alleles with a Frequency >= 5%, Ne = No. of Effective Alleles = 1 / (Sum pi^2), I = Shannon's Information Index = -1* Sum (pi * Ln (pi)), No. Private Alleles = No. of Alleles Unique to a Single Population, No. LComm Alleles (<=25%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 25% or Fewer Populations, No. LComm Alleles (<=50%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 50% or Fewer Populations, He = Expected Heterozygosity = 1 - Sum pi^2, and uHe = Unbiased Expected Heterozygosity = (2N / (2N-1)) * He

Allelic pattern across the populations revealed no locally common alleles were observed in 25% or 50 % of the populations of the tetraploid wheat study panel and the landraces based on DArTSeq markers. Private alleles are important in identifying unique genetic variability at loci and diverse genotypes that can be used in a breeding program to improve the allele richness in a study panel or population (Salem and Sallam 2016). In our current study, DArTSeq markers revealed 127 genotypes (126 landraces and only one released variety) to have one or more private allelels at 755 loci in the study panel in population one (data not shown), which is comprised of 83.15 % (153) of the study panel (150 landraces and 3 released varieties); however, no private allele was observed in population two, which is comprised of ninteen released varieties along with twelve released varieties. This could be an indication of the existance of key adaptive genes at these loci (Kumar *et al*., 2019). Allelic pattern revealed no locally common alleles were observed in 25% or 50 % populations of the tetraploid wheat study panel and the landraces based on DArTSeq markers. In line with this study, Sajjad *et al.*, (2018) reported that allelic patterns showed no locally common alleles in any of the subpopulation in Spring wheat populations by SSR-markers. The number of different alleles in population one and population two of the study panel were 4762 and 4708, respecively and 54 effective alleles were reported in population one of the study panel. The separate analysis of the landraces revealed the total number of alleles and the the total number of effective alleles in population one and population two of the landraces were 4624 and 4760 and 2 and 128, respectively. The mean value of number of different alleles and number of effective alleles were 2.000 ± 0.000, 1.423 ± 0.007 and 1.977 ± 0.003, 1.612 ± 0.006 respectively for population one and population two in the study panel. In line with this study, Kumar *et al*. (2019) reported the

mean value of number of different alleles and number of effective alleles for the two subpopulations were 2.0 and 1.834, respectively in Indian spring wheat. The mean values of Shannon's Information Index, based on DArTSeq markers, were 0.415 ± 0.004, and 0.523 ± 0.004 in population one and population two respectively for the tetraploid wheat study panel. The average Shannon's Information Index values for the landraces were 0.367 ± 0.005, 0.432 ± 0.004 respectively in population one and population two. The Shannon's Information Index (0.643, 0.648) reported by Kumar *et al*. (2019) in Indian spring wheat is a little bit higher than the present study result.

Table 5. 6 Total and mean allelic patterns across populations of the study panel and the landraces based on DArTSeq markers

	Mean allelic patterns across populations				**Total Allelic Patterns by Populations**			
	Study panel		**Landraces**		**Study panel**		**Landraces**	
Population	**Pop1**	**Pop2**	**Pop1**	**Pop2**	**Pop1**	**Pop2**	**Pop1**	**Pop2**
Na	2.000 ± 0.000	1.977 ± 0.003	1.942 ± 0.005	1.999 ± 0.001	4762	4708	4624	4760
Na Freq. >= 5%	1.858 ± 0.007	1.921 ± 0.006	1.677 ± 0.010	1.903 ± 0.006	4424	4574	3992	4532
Ne	1.423 ± 0.007	1.612 ± 0.006	1.380 ± 0.007	1.442 ± 0.006	54	0	2	138
I	0.415 ± 0.004	0.523 ± 0.004	0.367± 0.005	0.432 ± 0.004	0	0	0	0
No. Private Alleles	0.023 ±0.003	0.000 ± 0.000	0.001 ± 0.001	0.058 ± 0.005	0	0	0	0
No. LComm Alleles (<=25%)	0.000 ±0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000				
No. LComm Alleles (<=50%)	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000				
He	0.262 ±0.003	0.353 ± 0.003	0.234 ± 0.004	0.276 ± 0.003				
uHe	0.264 ±0.003	0.359 ± ±0.003	0.236 ± 0.004	0.277 ± 0.003				

Remarks: Na = No. of Different Alleles, Na (Freq >= 5%) = No. of Different Alleles with a Frequency >= 5%, Ne = No. of Effective Alleles = 1 / (Sum pi^2), I = Shannon's Information Index = -1* Sum (pi * Ln (pi)), No. Private Alleles = No. of Alleles Unique to a Single Population, No. LComm Alleles (<=25%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 25% or Fewer Populations, No. LComm Alleles (<=50%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 50% or Fewer Populations, He = Expected Heterozygosity = 1 - Sum pi^2, and uHe = Unbiased Expected Heterozygosity = (2N / (2N-1)) * He. Pop1 and pop2 respectively stands for population one and population two.

The mean values of Shannon's Information Index values were 0.415 ± 0.004, and 0.523 ± 0.004 in population one and population two respectively for the tetraploid wheat study panel. The average Shannon's Information Index values for the landraces were 0.367 ± 0.005, 0.432 ± 0.004 respectively in population one and population two.

The mean values of the number of different alleles and number of effective alleles were 2.000 ± 0.000, 1.423 ± 0.007 and 1.977 ± 0.003, 1.612 ± 0.006 respectively for population one and population two in the study panel. The mean values of number of different alleles and number of effective alleles for the two sub-populations (pop1-1, pop1-2) were respectively 1.942 ± 0.005 and 1.380 ± 0.007 and 1.999 ± 0.001 and 1.442 ± 0.006 in the landraces.

The mean values of expected heterozygosity (He) were 0.262 ±0.003 and 0.353 ± 0.003 respectively for population one and population two of the tetraploid wheat study panel and were 0.234 ± 0.004 and 0.276 ± 0.003 respectively for population one and population two of the tetraploid wheat landraces (Table 5.6).

The mean values of unbiased expected heterozygosity (uHe) were 0.264 ±0.003 and 0.359 ± ±0.003 respectively for population one and population two of the tetraploid wheat study panel and were 0.236 ± 0.004 and 0.277 ± 0.003 respectively for population one and population two of the tetraploid wheat landraces (Table 5.6). The total number of different alleles reported in the current study was 4762 in population one and 4708 in population two of the study panel and 4624 in population one and 4760 in population two of landraces. The total number effective alleles reported was 54 in population of one of the study panel, 2 in population one of the landraces and 128 in population of the landraces.

Allelic patterns across the species, regions of collections, and altitudinal classes of tetraploid wheat based on DArTSeq for the study panel were given in Figure 5.11 and Table 5.7.

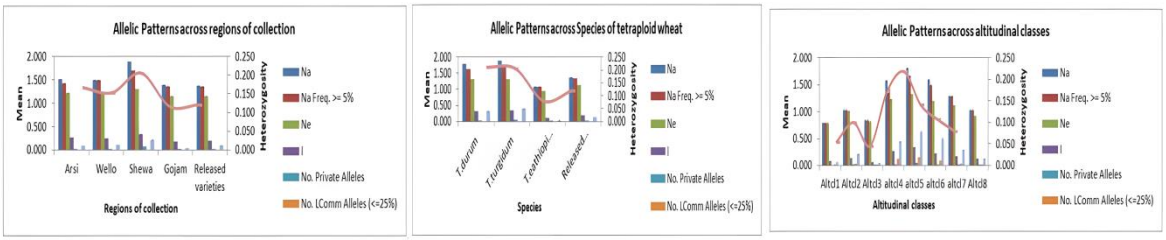


Figure 5. 11 Allelic patterns across the populations of tetraploid wheat germplasm at regions of collection, species level, and altitudinal classes based on DArTSeq markers. Remarks: Na = No. of Different Alleles, Na (Freq >= 5%) = No. of Different Alleles with a Frequency >= 5%, Ne = No. of Effective Alleles = 1 / (Sum pi^2), I = Shannon's Information Index = -1* Sum (pi * Ln (pi)), No. Private Alleles = No. of Alleles Unique to a Single Population, No. LComm Alleles (<=25%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 25% or Fewer Populations, No. LComm Alleles (<=50%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 50% or Fewer Populations, He = Expected Heterozygosity = 1 - Sum pi^2, and uHe = Unbiased Expected Heterozygosity = (2N / (2N-1)) * He

Allelic pattern across the populations revealed no locally common alleles were observed in 25% of the populations of regions of collections and species based on DArTSeq markers. The numbers of locally common alleles observed in the 50% of the populations were 241, 274, 490, 116, and 221 at the regions of collections and the species level respectively. The number of different alleles in population of the regions of collections were respectively 3583, 3544, 4511, 3293, and 3248 for Arsi, Welle, Shewa, Gojam, and Released varieties and the number of effective alleles were respectively 17, 25, 184, 10, and 36. The separate analysis of the species revealed that the total number of alleles and the total number of effective alleles in populations of Arsi, Wello, Shewa, Gojam, and released varieties were 4234, 4512, 2538, and 3248 and 90, 160, 4,and 36, repspectively. At the regions of collections, the mean private alleles was 0.007±0.002, 0.010±0.002, 0.077±0.005, 0.004±0.001, and 0.015±0.003 in the populations of Arsi, Wello, Shewa, Gojam, Released varieties, respectively. At the species level, the mean private alleles were 0.038±0.004, 0.067±0.005, 0.002±0.001, and 0.015±0.003 respectively for *Triticum durum*, *Triticum turgidum, Triticum aethiopicum*, and the released varieties. The altitudinal level analysis revealed that the number of effective alleles, the number of locally common alleles in 25% and in 50% of the populations were relatively higher in altitudinal classes 4, 5, and 6 when compared with other altitudinal classes.

The mean values of Shannon's Information Index in their decreasing order for the populations of regions of collections were 0.335±0.004, 0.258±0.005, 0.240±0.00, 0.189±0.005, and 0.184±0.005 for Shewa, Arsi, Wello, released varieties, and Gojam

respectively. At the species level, the mean Shannon's Information Index was 0.335±0.004, 0.332±0.005, 0.189±0.005, and 0.123±0.005 respectively for *Triticum turgidum, Triticum durum,* released varieties, and *Triticum aethiopicum in decreasing order*.

The respective mean values of number of different alleles and number of effective alleles were 1.505±0/012, 1.213±0.009, 1.488±0.012, 1.191±0.009, 1.895±0.006, 1.308±0.005, 1.383±0.012, 1.143±0.007, 1.364±0.012 and 1.141±0.008 respectively for Arsi, Wello, Shewa, Gojam, and released varieties in the populations of regions of collections. At the species level, the respective mean values of number of different alleles and number of effective alleles were 1.778±0.010, 1.313±0.008, 1.895±0.006, 1.306±0.005, 1.066±0.014, 0.936±0.011, 1.364±0.012, and 1.141±0.008 respectively for *Triticum durum, Triticum turgidum, Triticum aethiopicum,* and released varieties. At the altitudinal classes, the respective mean values of number of different alleles and number of effective alleles were 0.792±0.013, 0.778±0.012, 1.029±0.0013, 1.004±0.012, 0.840±0.012, 0.810±0.011, 1.572±0.012, 1.222±0.009, 1.804±0.009, 1.332±0.007, 1.596±0.010, 1.202±0.006, 1.279±0.012, 1.104±0.008, 1.030±0.014, and 0.924±0.011 respectively for class1, class 2, class3, class 4, class 4, class 5, class 6, class 7, and class8.

Table 5. 7 Total and mean allelic patterns across the species, regions of collections, and altitudinal classes of tetraploid wheat of Ethiopia based on DArTSeq markers

	Mean allelic patterns across regions of collections					Total allelic patterns across regions of collections				
Population	**Arsi**	**Wello**	**Shewa**	**Gojam**	**RV**	**Arsi**	**Wello**	**Shewa**	**Gojam**	**RV**
Na	1.505±0.012	1.488±0.012	1.895±0.006	1.383±0.012	1.364±0.012	3583	3544	4511	3293	3248
Na Freq. >= 5%	1.419±0.012	1.488±0.012	1.708±0.009	1.362±0.011	1.349±0.012	3378	3544	4066	3243	3213
Ne	1.213±0.009	1.191±0.009	1.308±0.005	1.143±0.007	1.141±0.008	17	25	184	10	36
I	0.258±0.005	0.240±0.005	0.335±0.004	0.184±0.005	0.189±0.005					
No. Private Alleles	0.007±0.002	0.010±0.002	0.077±0.005	0.004±0.001	0.015±0.003					
No. LComm Alleles (<=25%)	0.000±0.000	0.000±0.000	0.000±0.000	0.000±0.000	0.000±0.000	0	0	0	0	0
No. LComm Alleles (<=50%)	0.101±0.006	0.115±0.007	0.206±0.008	0.049±0.004	0.093±0.006	241	274	490	116	221
He	0.167±0.004	0.153±0.004	0.206±0.003	0.116±0.003	0.121±0.003					
uHe	0.181±0.004	0.167±0.004	0.213±0.003	0.127±0.004	0.131±0.004					

	Mean allelic patterns across species				Total allelic patterns across species			
Population	**T.durum**	**T.turgidum**	**T.eathiopicum**	RV	**T.durum**	**T.turgidum**	**T.aethiopicum**	**RV**
Na	1.778±0.010	1.895±0.006	1.066±0.014	1.364±0.012	4234	4512	2538	3248
Na Freq. >= 5%	1.635±0.011	1.723±0.009	1.066±0.014	1.349±0.012	3892	4102	2538	3213
Ne	1.313±0.008	1.306±0.005	0.936±0.011	1.141±0.008	90	160	4	36
I	0.332±0.005	0.335±0.004	0.123±0.005	0.189±0.005				
No. Private Alleles	0.038±0.004	0.067±0.005	0.002±0.001	0.015±0.003				
No. LComm Alleles (<=25%)	0.000±0.000	0.000±0.000	0.000±0.000	0.000±0.000	0	0	0	0
No. LComm Alleles (<=50%)	0.318±0.010	0.393±0.011	0.028±0.003	0.145±0.008	758	935	67	345
He	0.211±0.004	0.206±0.003	0.079±0.003	0.121±0.003				
uHe	0.226±0.004	0.213±0.003	0.087±0.003	0.131±0.004				

	Mean allelic patterns across altitudinal classes							
Population	**Altcl1**	**Altcl2**	**Altcl3**	**altcl4**	**altcl5**	**altcl6**	**altcl7**	**Altcl8**
Na	0.792±0.013	1.029±0.0013	0.840±0.012	1.572±0.012	1.804±0.009	1.596±0.010	1.279±0.012	1.030±0.014
Na Freq. >= 5%	0.792±0.013	1.029±0.013	0.840±0.012	1.442±0.012	1.659±0.010	1.494±0.011	1.279±0.012	1.030±0.014
Ne	0.778±0.012	1.004±0.012	0.810±0.011	1.222±0.009	1.332±0.007	1.202±0.006	1.104±0.008	0.924±0.011

I	0.078±0.004	0.141±0.006	0.067±0.004	0.269±0.005	0.342±0.005	0.229±0.005	0.166±0.005	0.122±0.005
No. Private Alleles	0.000±0.000	0.006±0.002	0.001±0.001	0.025±0.003	0.055±0.005	0.026±0.003	0.009±0.002	0.002±0.001
No. LComm Alleles (<=25%)	0.012±0.002	0.039±0.004	0.005±0.001	0.116±0.007	0.158±0.008	0.098±0.006	0.043±0.004	0.015±0.002
No. LComm Alleles (<=50%)	0.070±0.005	0.203±0.008	0.045±0.004	0.441±0.011	0.622±0.013	0.493±0.012	0.280±0.010	0.120±0.007
He	0.055±0.003	0.100±0.004	0.046±0.003	0.172±0.004	0.218±0.003	0.141±0.003	0.107±0.003	0.080±0.003
uHe	0.070±0.004	0.126±0.005	0.055±0.003	0.185±0.004	0.233±0.004	0.150±0.003	0.119±0.004	0.090±0.004

Total allelic patterns across altitudinal classes

Population	**Altcl1**	**Altcl2**	**Altcl3**	**altcl4**	**altcl5**	**altcl6**	**altcl7**	**Altcl8**
Na	1886	2449	2001	3742	4296	3799	3045	2453
Na Freq. >= 5%	1886	2449	2001	3433	3950	3558	3045	2453
Ne	0	15	3	59	131	63	22	5
No. LComm Alleles (<=25%)	29	92	12	277	376	234	102	35
No. LComm Alleles (<=50%)	167	483	108	1051	1481	1173	666	285

Remark: RV=released varieties and altcl = altitudinal classes

The mean values of expected heterozygosity (He) were 0.167±0.004, 0.153±0.004, 0.206±0.003, 0.116±0.003, and 0.121±0.003 respectively for Arsi, Wello, Shewa, Gojam, and released varieties at regions of collections; 0.211±0.004, 0.206±0.003, 0.079±0.003, and 0.121±0.003 respectively for *Triticum durum, Triticum turgidum, Triticum aethiopicum* and released varieties at the species level, and 0.055±0.003, 0.100±0.004, 0.046±0.003, 0.172±0.004, 0.172±0.004, 0.218±0.003, 0.141±0.003, 0.107±0.003, and 0.080±0.003 respectively for class1, class2, class3, class4, class5, class6, class7, and class8 at the altitudinal level (Table 5.7).

The mean values of unbiased expected heterozygosity (uHe) were 0.181±0.004, 0.167±0.004, 0.213±0.003, 0.127±0.004, and 0.131±0.004 respectively for Arsi Wello, Shewa, Gojam, and released varieties at the regions of collection; 0.226±0.004, 0.213±0.003, 0.087±0.003, and 0.131±0.004 respectively for *Triticum durum, Triticum turgidum, Triticum aethiopicum,* and released varieties at the species level, and 0.070±0.004, 0.126±0.005, 0.055±0.003, 0.185±0.004, 0.233±0.004, 0.150±0.003, 0.119±0.004, and 0.090±0.004 respectively for class1, class2, class3, class4, class5, class6, class7, and class8.

5.3.6. Analysis of molecular variance

Analysis of molecular variance, based on DArTSeq markers, was conducted for tetraploid wheat panel, the landraces, regions of collections, species level, and altitudinal classes (Table 5.8). AMOVA revealed highly significant ($P<0.001$) variation within and among the populations of tetraploid wheat genotypes, the landraces, regions of collections, species level, and altitudinal classes. The AMOVA of the tetraploid wheat panel, the landraces, regions of collections, species level, and altitudinal classes revealed 12%, 3%, 16%, 16%, and 3% of the total variation was due to between populations and 88%, 97%, 84%,84%, and 97% of the total variation was due to within populations, respectively. There was more variation within populations than between populations. In line with this study, Negisho *et al*. (2021) reported higher percentage of variation within populations than between populations of durum wheat study panel and Kumar *et al*. (2019) reported 2% variation among subpopulations and 98% of variation within subpopulations in Indian spring wheat. The justification for high variation within populations is the selection for several agronomically important traits in various breeding programs (Kumar *et al*., 2019).

Fixation index (Fst) and gene flow (Nm) for the study panel, the landraces, regions of collections, species level, and altitudinal classes were Fst = 0.117 and Nm = 1.891, Fst = 0.028 and Nm = 8.612, Fst= 0.162, Nm= 1.289, Fst= 0.162 and Nm= 1.297, and Fst = 0.028 and Nm=8.818 respectively (Table 5.8). In line with the present study, Negisho *et al*. (2021) reported Fst = 0.19 and Nm = 1.04 in Ethiopian durum wheat study panel. Fixation index is a useful index of genetic differentiation (Hartl and Clark 1997). Fixation index (Fst), which is a measure of population differentiation due to genetic structure, can be significant when its value is greater than 0.15 (Negisho *et al*., 2021). Wright (1978) suggested the following qualitative guidelines for the interpretation of Fst: the range 0 to 0.05 may be considered as indicating little genetic differentiation, the range 0.05 to 0.15 indicates moderate genetic differentiation, the range 0.15 to 0.25 indicates great genetic differentiation, and values of Fst above 0.25 indicate very great genetic differentiation. Wright also suggested that among sub populations, differentiation is by no means negligible if Fst is as small as 0.05 or even less (Hartl and Clark 2007). Accordingly, DArTSeq markers revealed little (Fst < 0.05) genetic differentiation in the landraces (Fst=0.028) and in the altitudinal classes (Fst=0.028), moderate (Fst= 0.05 to 0.15) genetic differentiation in the study panel (Fst=0.117), and great genetic differentiation (Fst=0.15 to 0.25) in the regions of collections (Fst=0.162) and at the species level (Fst=0.162). The Fst values between populations of the landraces (0.028) and tetraploid wheat study panel (0.117) respectively indicates little to moderate genetic differentiation between these populations. This coincides with the result obtained from AMOVA, in which low percentage of the total variation was accounted for by among-subpopulation variations. The observed high variation within populations is an implication for more frequent selection for economically important traits (Kumar *et al*. 2019). DArTSeq markers revealed high gene flow in the case of the landraces as compared to the study panel. According to Wright (1978), gene flow (Nm) is divided into three grades: high (≥1.0), medium (0.250–0.99), and low (0.0–0.249). Accordingly, high gene flow was observed for the populations of tetraploid wheat at study panel, landraces, regions of collections, species level, and altitudinal classes. In the present study, there is more gene flow in the populations of tetraploid wheat landraces as compared to the population of tetraploid wheat study panel. The higher gene flow observed in the present study resulted in the loss of genetic differentiation among subpopulations (Kumar *et al*., 2019). Therefore, the Fst values for the tetraploid wheat study panel (0.117) and the landraces (0.028) indicated that there is no significant differentiation between the populations based on DArTSeq markers.

Eventually, the lower genetic differentiation led to more gene flow (Nm) values of Nm = 1.891 and Nm = 8.612 for the tetraploid wheat study panel and the landraces, respectively (Table 5.8). Gene flow (Nm) values greater than one is an indication of more gene exchange (Hartl and Clark 1997; Frankham *et al.*, 2002) preventing genetic differentiation (Frankham *et al.*, 2002). High gene flow value is an indication of a low level of diversity among subpopulation (Kumar *et al*., 2019) corroborated by AMOVA result in both the study panel and the landraces. The gene flow values based on the DArTSeq markers (Nm= 1.891) for the whole tetraploid wheat study panel in our study is indicating that on average two migrants per generation are entering the tetraploid wheat populations. In the case of tetraploid wheat landraces, Nm= 8.612 is an indication of the entrance of 9 migrants per generation in tetraploid wheat landraces. Migration is potent force acting against genetic divergence among subpopulations and hence, little migration is enough to prevent significant genetic differentiation among subpopulations (Hartl and Clark, 2007). Opposite to this study, Zhang *et al*. (2011) reported the estimated fixation index (Fst) value of 0.28 in common wheat cultivars in China indicating the germplasm is highly differentiated. Kumar *et al*. (2019) reported 0.196 and 0.532 values of Fst in sp1 and sp2, respectively in Indian spring wheat populations. The pinpointing of the causes underlying a particular value of Fst observed in a natural population is often very difficult; however, this doesn't invalidate the importance of Fst as an index of genetic differentiation. Differences in allele frequency among subpopulations could occur due to random genetic drift as well as natural selection complicated with migration among the subpopulations (Hartl and Clark 1997).

Table 5. 8 AMOVA of tetraploid wheat germplasm at study panel, landraces, regions of collections, species, and altitudinal class level

Summary AMOVA Table of DArTSeq of 184 tetraploid wheat genotypes						
Source of variation	df	SS	EV	PV	Fixation index	Gene flow
Among Pops	1	16994.550***	287.462	12%	0.117	1.891
Within Pops	182	395785.190***	2174.644	88%		
Total	183	412779.7	2462.106	100%		
Summary AMOVA Table Landraces based on DArTSeq markers						
Source	df	SS	EV	PV	PhiPT	Nm
Among Pops	3	36019.904***	167.219	3%	0.028	8.612
Within Pops	158	930682.615***	5890.396	97%		
Total	161	966702.5	6057.616	100%		
Summary AMOVA Table of species of tetraploid wheat genotypes based on DArTSeq markers						
Source	df	SS	EV	PV	PhiPT	Nm
Among Pops	3	38920.789***	290.667	16%	0.162	1.297
Within Pops	180	271451.363***	1508.063	84%		
Total	183	310372.2	1798.73	100%		
Summary AMOVA Table of regions of collection of tetraploid wheat genotypes						
Source	df	SS	EV	PV	PhiPT	Nm
Among Pops	4	41664.113***	291.072	16%	0.162	1.289
Within Pops	179	268708.039***	1501.162	84%		
Total	183	310372.2	1792.234	100%		
Summary AMOVA Table altitudinal classes of landraces of tetraploid wheat genotypes						
Source	df	SS	EV	PV	PhiPT	Nm
Among Pops	7	15479.053***	41.712	3%	0.028	8.818
Within Pops	154	226580.083***	1471.299	97%		
Total	161	242059.1	1513.011	100%		

Remarks: Probability, P (rand >= data), for PhiPT is based on standard permutation across the full data set. PhiPT = AP / (WP + AP) = AP / TOT, Nm = [(1 / PhiPT) - 1] / 4, and key: AP = Estimated Variance Among Populations, WP = Estimated Variance Within Populations, EV= Estimated Variance, PV= Percentage of variation.

5.3.6. Cluster analysis

Cluster analysis, based on population structure, for the whole tetraploid wheat study panel and the landraces was given in Figure 5.4 and Figure 5.6 respectively and illustrated under the section population structure analysis. Cluster analysis for tetraploid wheat based on regions of collections, species level, and altitudinal classes (Figure 5.12) was generated using Neighbor Joining (NJ) tree from Euclidean distance based on Ward.D method using 2381 DArTSeq markers. Cluster analysis at the regions of collections revealed that genotypes from Shewa and released varieties were grouped together, genotypes from Wello remained solitary, and the genotypes from Gojam and Arsi were grouped together; at species level, *Triticum durum* and released varieties were grouped together, *Triticum aethiopicum* and *Triticum turgidum* were grouped together, and *Triticum polonicum* remained solitary and at altitudinal classes, only genotypes from altitudinal class 5 and class 6 were grouped together while others remained solitary. The grouping together of landraces from mutually distant regions of Arsi and Gojam could be due to the introduction of germplasm from one to the other, either formally or informally, sometimes long ago (Mengistu *et al*., 2015). The possible reason for the clustering of released varieties with the landarces from Shewa could be the recollection of improved varieties disseminated to the the farmers and under cultivation for longer periods and considered as landraces by farmers (Mengistu *et al*., 2015). The clustering patten using DArTSeq markers and grain quality traits showed similar trend in grouping released varieties with the landraces from Shewa; however, the clustering pattern based on phenotypic traits, grain quality traits, and DArSTeq markers did not show similarity except in the clustering pattern based on grain quality traits and DArTSeq markers, which grouped released varieties with landraces from Shewa. Similar result was reported by Mengistu et al. (2015) for landraces from SNNP and released varieties in Ethiopian durum wheat landarces.

The maximum (0.311, 0.226, and 0.062) Euclidean distance respectively was recorded between *Triticum polonicum* and released varieties at the species level, between released varieties and Arsi for the regions of collections, and between class 1 and class 2 at the altitudinal level; while, the minimum (0.007, 0.011, and 0.007) Euclidean distance was recorded between *Triticum durum* and *Triticum turgidum* at the species level, between Gojam and Wello for the regions of collections, and between class 4 and class 6 and class 5 and class 6 at the altitudinal level (Table 5.9). Similar to the present study, Hailu *et al.* (2005) reported small Euclidean distance between *Triticum turgidum* and *Triticum* durum

followed by the distance between *Triticum aethiopicum* and *Triticum durum* based on ISSR; however, opposite to this study, they reported large distance between *Triticum turgidum* and *Triticum aethiopicum* as compared to others. The maximum (0.387) and the minimum distances (0.011) in the study panel were observed between ETW115 (7683) and ETW135 ('MCD11-20, MCD8- 2, ETDW/15DZ019,……') and ETW104 (235051) and ETW105 (7507), respectively. The maximum (0.553) and minimum (0.0012) distance in the released cultivars, based on the DArTSeq markers, were respectively observed between ETW185 (Tesfaye) and ETW 169 (Bekelcha) and ETW185 (Tesfaye) and ETW187 (Alemtena) (Fufa *et al.,* 2023).

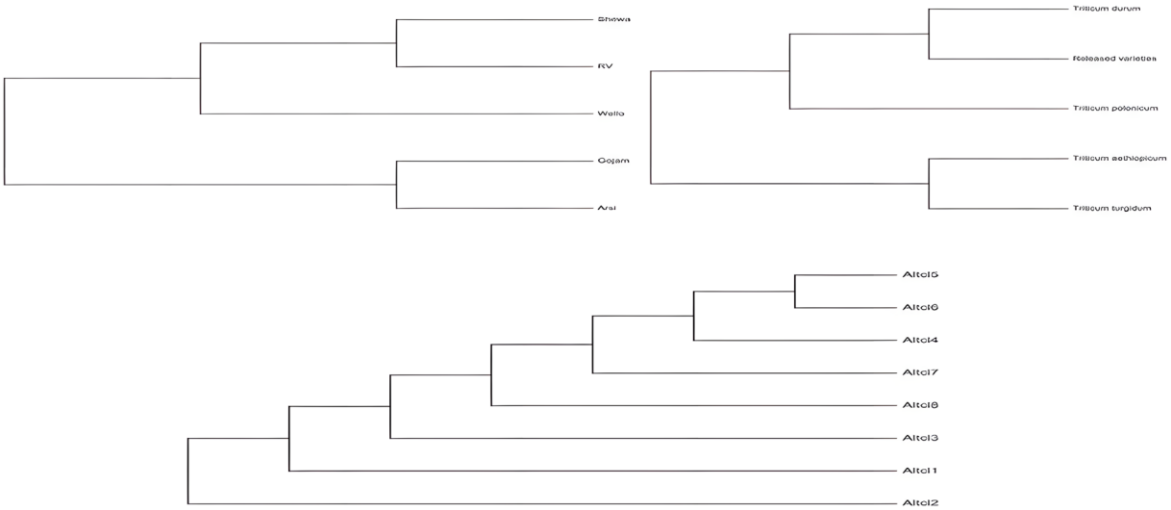


Figure 5. 12 Cluster analysis of tetraploid wheat genotypes based on regions of collections, species level, and altitudinal classes using DArTSeq markers. Remarks: RV = released varieites and Altcl1-8 represents altitudinal classes from one eight.

Table 5. 9 Euclidean distances among the populations of tetraploid wheat genotypes at species, regions of collections, and altitudinal classes

	Species					Regions of collections				
	RV	*TP*	*TE*	*TD*	*TR*		Arsi	Gojam	RV	Shewa	Wello
RV	0					Arsi	0				
TP	0.311	0				Gojam	0.025	0			
TE	0.232	0.166	0			RV	0.226	0.215	0		
TD	0.151	0.142	0.022	0		Shewa	0.035	0.015	0.162	0	
TR	0.188	0.139	0.017	0.007	0	Wello	0.037	0.011	0.192	0.012	0

	Altitudinal classes							
	Altcl1	Altcl2	Altcl3	Altcl4	Altcl5	Altcl6	Altcl7	Altcl8
Altcl1	0							
Altcl2	0.062	0						
Altcl3	0.052	0.056	0					
Altcl4	0.039	0.035	0.029	0				
Altcl5	0.027	0.034	0.03	0.01	0			
Altcl6	0.035	0.038	0.033	0.007	0.007	0		
Altcl7	0.039	0.044	0.042	0.016	0.01	0.014	0	
Altcl8	0.046	0.051	0.041	0.019	0.016	0.013	0.02	0

Remarks: RV= released varieties, Altcl= altitudinal classes, TP= *Triticum polonicum*, TE= *Triticum aestivum*, TD= *Triticum durum*, and TR= *Triticum turgidum*.

5.5. Conclusion

In the present study, genetic diversity and population structure analysis of 184 tetraploid wheat study panel was conducted using high quality DArTSeq markers. AMOVA ($p<0.001$) revealed that the genetic variation within the accessions of populations is higher than the genetic variation between populations. Population structure analysis of the study panel showed two separate populations (Pop1 and Pop2), which is also verified by the principal component analysis. Genetic diversity indices for the study panel revealed higher genetic diversity in Pop2, which comprised 19 released varieties along with 12 landraces. The genotypes with maximum (0.387, 0.553) Euclidean distances in the study panel and in the released cultivars, ETW115 (7683) and ETW135 ('MCD11-20, MCD8- 2, ETDW/15DZ019,……') and ETW185 (Tesfaye) and ETW 169 (Bekelcha), could be used in a crossing programme.

DArTSeq markers revealed little (Fst = 0.028) and moderate (Fst = 0.117) genetic differentiation in the landraces and in the study panel, respectively implying no significant differentiation between the populations. Eventually, the lower genetic differentiation led to more gene flow, which is an indication of more gene exchange (Hartl and Clark, 2007; Frankham *et al.*,2002) preventing genetic differentiation (Frankham *et al.*, 2002) among populations implying a low level of diversity among subpopulation (Kumar *et al*., 2019).

The present study revealed private alleles in population one, comprising 98.04 % landraces and 1.06 % released varieties, of the study panel at 755 loci implying the huge genetic potential held in the tetraploid wheat germplasm of Ethiopia and this could be an indication of the existance of key adaptive genes at these loci (Kumar *et al*., 2019). One hundred twenty-seven genotypes (126 landraces and one released variety) were reported to have one or more private alleles, which are important in identifying unique genetic variability at these loci and diverse genotypes that can be used in a breeding program to improve the allele richness in a study panel or population (Salem and Sallam 2016).

Therefore, the high genetic diversity detected in Ethiopian tetraploid wheat germplasm showed the existence of huge variability that could be utilized for future wheat breeding program. Furthermore, the higher genetic variation observed within the populations than between the populations of the genotypes reveal that more attention should be given to individual accessions within populations to explore the existing genetic diversity for genetic material conservation, improvement program, and as a basis for genomic analysis.

CHAPTER VI

6. Genome-Wide Association Study in Tetraploid Wheat (*Triticum turgidum* spp.) Germplasm of Ethiopia

Abstract

*Genome-wide association study (GWAS) helps the dissection of complex traits through the identification of DNA markers associated with traits of economic interest and thus, enhances the improvement of the productivity of crops. Tetraploid wheat species of Ethiopia are gifted with desirable alleles useful for economically important traits. We used a panel of 184 tetraploid wheat (Triticum turgidum spp) genotypes, including 162 landraces and 22 improved varieties, which were retained after filtering of genotypes genotyped using high density genotyping by sequencing (GBS) based on DArTSeq markers. A total of 2381 robust SNPs (0.35≤ PIC ≤ 0.50 and call rate ≥ 0.75) were used for GWAS analysis using **Bayesian**-information and Linkage-disequilibrium Iteratively Nested Keyway (**BLINK**) model based on the best linear unbiased estimate (BLUE) values produced from a combined data at Sinana and Debrezeit for two years. Analysis of genome-wide LD in the whole collection showed that 44.98% of the total marker pairs had a significant LD ($p<0.01$) with a mean r^2 value of 0.129791. LD decayed to its half at 7.68Mbp for the whole genome. The GWAS scan identified 44 significant MTAs, two of which are pleiotropic, across chromosome 1A, 2A, 2B, 3A, 3B, 4B, 5A, 5B, 6A, 6B, and 7B. Nine of these MTAs are novel, whereas the remaining 35 were previously reported and confirmed in this study. We also identified candidate genes for the novel loci potentially regulating the traits. Hence, this study highlights the significance of the Ethiopian tetraploid wheat gene pool for improving tetraploid wheat globally. Furthermore, a breeding strategy focusing on accumulating favorable alleles at these loci could improve tetraploid wheat production in Ethiopia and internationally.*

Keywords: BLINK; DArTSEq markers; GWAS; Tetraploid wheat

6.1. Introduction

Ethiopia, the largest wheat producer in sub-Saharan Africa, is a center of diversity for tetraploid wheat (Hodson *et al*., 2020; Haile *et al*., 2013), which are gifted with valuable wealth of genetic diversity with untapped primary sources of desirable genes useful for economically important traits (Muleta *et al*., 2017). The key to modern plant breeding is elucidating the genetic basis of complex quantitative traits controlling agronomic traits of economic importance (Alemu *et al*., 2020) by undertaking genome wide association mapping, which is a powerful technique used to identify marker-trait associations.

Genome-wide association (GWA) mapping, also called whole genome scanning approach, is a comprehensive approach in which genotyping is done for all possible genetic markers across the genome to systematically search the genome for causal genetic variation affecting the target trait(s) (Hirschhorn and Daly 2005), identifies genomic regions throughout the genome associated with the trait of interest (Remington *et al*., 2001), and is useful to facilitate crop improvement via enhanced knowledge of marker-trait associations (MTA) (Lozada *et al*., 2017). This approach is favoured in situations where linkage disequilibrium (LD) extends for large distances, allowing for potential candidate regions associated with a trait of interest to be identified for further study (Remington *et al*., 2001). GWAS has the potential to identify a single polymorphism within a gene that is responsible for phenotypic differences. Association Mapping involves searching for genotype-phenotype correlations among unrelated individuals and has high resolution that is accounted for by the historical recombination accumulated in natural populations and collections of landraces, breeding materials, and varieties (Mahmut Caliskan 2012).

Undertaking association mapping is of great value to discern associations between genetic markers and traits of interest (Hu *et al*., 2015; Mwadzingeni *et al*., 2017) and has been widely used to dissect the genetic architecture behind traits such as grain yield and yield related traits (Gao *et al*., 2021; Bhatta *et al*., 2018; Alemu *et al*., 2021; Jung *et al*., 2021; Mengistu *et al*., 2016; Kidane *et al*., 2019; Mérida-García *et al*., 2019), host plant resistance to pathogens (Maccaferri *et al*. 2010; Aoun *et al*. 2016; Negash *et al*. 2022), drought tolerance (Bhatta *et al*., 2018; Mathew *et al*., 2019; Alemu *et al*., 2021; Turki *et al*., 2014), root architecture (Cane *et al*., 2014), and quality traits (Alemu *et al*., 2021; Gao *et al*., 2021; Maccaferri et al. 2010; Tuberosa 2012; Chen *et al*., 2019; Mérida-García *et al*., 2019; Giraldo *et al*., 2016;Talini *et al*., 2020) in wheat. GWAS results in tetraploid

wheat of Ethiopia were reported on the following traits of economic importance including yield and yield related traits (Kidane *et al*., 2017; Mengistu *et al*., 2016), root system architecture (Alemu *et al*., 2021), disease resistance (Letta *et al*., 2014; Kidane *et al*., 2017; Alemui *et al*., 2021), grain shape and color (Alemu *et al*., 2020). The findings so far done on GWAS in tetraploid wheat of Ethiopia revealed important insight on the association of traits of economic importance and genetic markers; however, it is limited and insufficient. Tetraploid wheat species of Ethiopia are endowed with valuable wealth of genetic diversity with untapped primary sources of desirable genes useful for economically important traits (Haile *et al*., 2013; Muleta *et al*., 2017). Thus, understanding the genetic basis of these important traits using recent genomic based research will facilitate the utilization of Ethiopian tetraploid germplasm in the crop improvement program. Diversity Arrays Technology (DArT) can detect any type of DNA variation at several hundred genomic loci in parallel without relying on sequence information (Wenzl *et al*., 2016) and has been widely and successfully applied for polyploidy and complex genomes such as tetraploid wheat (Robbana *et al*., 2019) for association mapping studies in many plant species (Ficco *et al*., 2012) such as wheat (Akbari *et al*., 2006); however, association mapping using DArT markers in tetraploid wheat of Ethiopia is not available. The application of DArT markers in the association mapping in tetraploid wheat of Ethiopia would reveal valuable alleles of economic importance. Therefore, this study was initiated to detect the association of phenotypic traits and grain quality traits with DArTSeq markers.

6.2. Materials and Methods

A panel of 184 tetraploid wheat (*Triticum turgidum spp*) genotypes, including 162 landraces and 22 improved varieties, which were retained after filtering of genotypes genotyped using high density genotyping by sequencing (GBS) by Diversity Arrays Technology sequencing were used for the Genome Wide Association Study (GWAS) (Annex 1). The association mapping of multiple phenotypic traits and grain quality traits with genome-wide scanned SNPs was carried out using the Genome Association and Prediction Integrated Tools (GAPIT) package (Lipka *et al*., 2012) in the R software (R Development Core Team 2018) based on the best linear unbiased estimate (BLUE) values produced from two-year combined data at Sinana and Debrezeit during 2020 and 2021. The analysis involved a total of 2381 ($0.35 \leq$ PIC ≤ 0.50 and call rate ≥ 0.75) robust SNPs. Missing data were imputed using optimal impute ver. 1.0.5 in the KDcompute_plugin system based on the KNN imputation method. The marker distribution on each chromosome was determined using LD measure in R^2 ver.0.2.2 of the KDcompute_plugin. Pairwise LD measures (r^2 and P-value) between markers on each chromosome were also computed using TASSEL Ver.5 (Bradbury *et al*., 2007). A genome-wide LD decay scatter plot was then produced by plotting the r^2 values against physical distance (bp) using the GAPIT software. Finally, r^2 =0.2 was considered as a cut-off point for no LD between pairs of markers.

The specific critical r^2 value beyond which LD is due to true physical linkage was determined by taking the 95^{th} percentile of r^2 data of unlinked marker pairs (Breseghello and Sorrells 2006). LD decay was estimated as physical distance using the procedure of Hill and Weir (Hill and Weir 1988) in R (R Development Core Team 2018) and an LD decay curve was fitted with a smoothing spline regression line at the genome level. The genome-wide association (GWAS) analysis was conducted using Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) model (Huang *et al*., 2019) implemented in the GAPIT R package (R Development Core Team 2018) with population structure (Q) and kinship (K) covariates. The algorism uses both fixed-effect and random-effect models iteratively to control spurious marker-trait association due to population structure and family relatedness (Lipka *et al*., 2012). Further, a kinship (K) matrix was computed using the method of Vanraden (2008). Principal components analyse describing the population stratification were computed using R/GAPIT and iteratively added to the fixed part of the model. Quantile-quantile (Q-Q) plots generated from –

log10p-values were assessed visually to determine how well the model accounted for population structure and family relatedness among the study samples. Statistically significant marker-trait associations were declared using a false discovery rate (FDR)-adjusted $p < 0.05$ as implemented in GAPIT. Furthermore, the Bonferroni correction rate at a significance threshold of $p < 0.05$ or –log10 (p-values) = 3 was also included in the analysis for comparison. Q-Q and Manhattan plots were visualized using the R package qqman (Turner 2018). The high-confidence candidate genes within the identified phenotypic and grain quality traits associated regions were also extracted from the recently released IWGSC RefSeq Annotation v2.1 available on the URGI Seq repository (https://wheat-urgi.versailles.inrae.fr/Seq-Repository/Annotations). Significance intervals of identified QTLs were reported as the intervals after inclusion of all SNPs associated with the trait with $P \leq 0.01$ (marker-wise) and in LD of $r^2 \geq 0.3$. The confidence interval of a QTL was defined based on ±3.84 Mbp (LD decay of this panel) from the GWAS QTL peak (QTL-tag SNP).

6.3. Results and Discussion

For GWAS 184 study panel of tetraploid accessions, including 22 improved varieties and 162 landraces collected from major wheat-producing areas of the of Ethiopia including Bale, Arsi, Shewa, Jimma, Agaw, Gonder, Gojam, Tigray, and Wello (Figure 3.1) were used (Annex1). Genome wide association study (GWAS) was carried out based on the best linear unbiased estimate (BLUE) values obtained from the combined analysis of two years data at Sinana and Debrezeit. The analysis involved a total of 2381 ($0.35 \leq$ PIC ≤ 0.50 and call rate ≥ 0.75) robust SNPs.

6.3.1. Combined ANOVA: Combined ANOVA from two years data produced from the evaluation of 196 genotypes, including 22 improved varieties and 174 landraces of tetraploid wheat at Sinana and Debrezeit is given in Table 6.1. The combined analysis of variance revealed highly significant variation among genotypes for all traits.

Table 6. 1 Mean squares from combined ANOVA of phenotypic and grain quality traits showing significant MTA in tetraploid wheat of Ethiopia

SV	df	SdPSp	PH	DB	DM	TKW	GY	VITP	ZI	HI
Rep	1	0.8006*	2683***	189**	628***	514***	0.2	320	461*	1005**
Gen	195	1.751***	331***	78***	37***	89***	3.3***	1949***	384***	284***
Loc	1	1.78**	111410***	9805***	238050***	1053***	0.2	21090***	13699***	21432***
Yr	1	0.5787	88372***	4461***	1331***	325***	97.4***	3236***	38835***	3461***
GenxLoc	195	0.3459***	81	27**	31**	30***	2***	316***	102**	210***
GenxYr	195	0.0673	69	22	26	28**	1.6***	450***	113***	269***
LocxYr	1	0.0249	187364***	1	1034***	30705***	400.2***	31659***	4300***	
Rep:block	26	0.1468	82	37**	25	21	1	331*	111	223
Gen:Loc:Yr	195	0.0549	92	27**	30*	34***	2.7***	940***	143***	
Residuals	757	0.1724	71	21	23	21	1.1	212	79	97
CV		16	17.6	8	11.6	20.1	39	30.1	25.2	45.8
Mean		2.54	107.37	71.76	122.43	34.65	3.51	66.93	46.6	29.25

Remark: SV= sources of variation, df = degrees of freedom, Rep=replication, Gen=genotypes, Loc= location, Yr=year, CV = coefficient of variation, SdPSp=number of seeds per spikelets, PH=plant height, DB= days to booting, DM= days to maturity, TKW =thousand kernel weight, GY= grain yield, VITP= vitreousness percentage, ZI=Zeleny index, and HI= harvest index.

6.3.2. Frequency distribution of traits having significant MTAs: Normal distribution was observed for days to booting, days to maturity, vitreousness percentage, zeleny index, grain yield, plant height, thousand kernel weight, number of seeds per spikelets; however, white seed color, dense spike density, very lux spike density, and white glume color showed bimodal distributions (Figure 6.1).

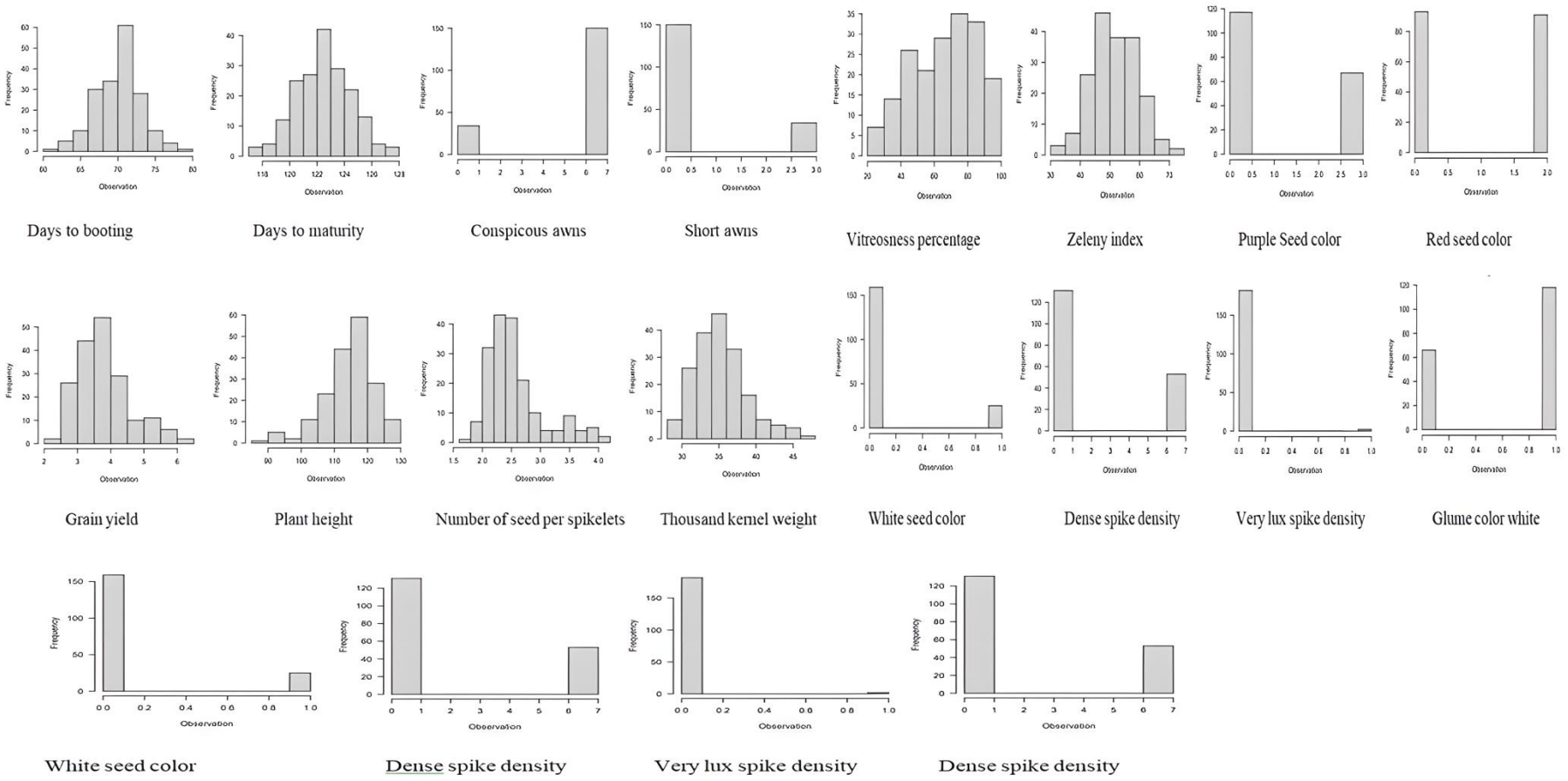


Figure 6. 1 Frequency distribution of traits having significant marker trait association in 184 tetraploid wheat germplasm of Ethiopia

6.3.3. Linkage disequilibrium: Significance intervals of identified QTL were reported as the intervals after inclusion of all SNPs associated with the trait with $P \leq 0.01$ (marker-wise) and in LD of $r^2 \geq 0.3$. The confidence interval of a QTL was defined based on ±3.84 Mbp (LD decay of this panel) from the GWAS QTL peak (QTL-tag SNP).

Table 6. 2 Summary of linkage disequilibrium analyses among marker pairs and the number of significant marker pairs per chromosome and genome

Chromosome	TNMP	Distance (Mbp)	r^2	Significant markers ($p < 0.01$)
1A	5659	112.2715	0.168363	3059(3.29)
1B	7054	105.5453	0.13045	3199(3.44)
2A	7359	111.3006	0.137588	2670(2.87)
3A	5746	136.3984	0.131523	1719(1.85)
4A	3975	169.2413	0.127932	2577(2.77)
4B	4421	160.8146	0.15817	1675(1.80)
3B	7863	118.2308	0.215542	4218(4.53)
2B	9378	89.26415	0.103171	3762(4.04)
5A	6506	99.07598	0.1199	3865(4.15)
5B	7803	93.3874	0.115398	4568(4.91)
6A	4262	138.9247	0.117751	2173(2.33)
6B	5588	127.8263	0.114819	3673(3.95)
7A	10417	78.71392	0.111175	2231(2.40)
7B	7045	113.6408	0.083326	2477(2.66)
A genome	43924	113.0938	0.129077	19117(20.54)
B genome	49152	111.2032	0.13043	22749(24.44)
A+B Genome	93076	112.0952	0.129791	41866(44.98)

The LD decayed to its half at 7.68Mbp for the whole genome at r^2 value of 0.231 (Figure 6.2), which is moderate level of LD. In line with this study, Kidane *et al.* (2019) reported an LD decay distance of 7.4 Mb on average in durum wheat of Erhiopia and Mulugeta *et al.* (2023) reported $r^2 = 0.21$ in durum wheat of Ethiopia. Analysis of genome-wide LD in the whole accessions showed that 44.98% of the total marker pairs had a significant LD ($p<0.01$) with a mean r^2 value of 0.129791. 44.59% (18668) and 8.62% (3609) of the significant marker pairs showed r^2 values above 0.2 and 0.7, respectively, indicating higher level of LD in the genome. The number of significant marker pairs observed was higher in B genome (24.44%) than A genome (20.54%) and the highest and lowest number was recorded in 5B (4.91%) and 4B (1.80%), respectively (Figure 6.3; Table 6.2).

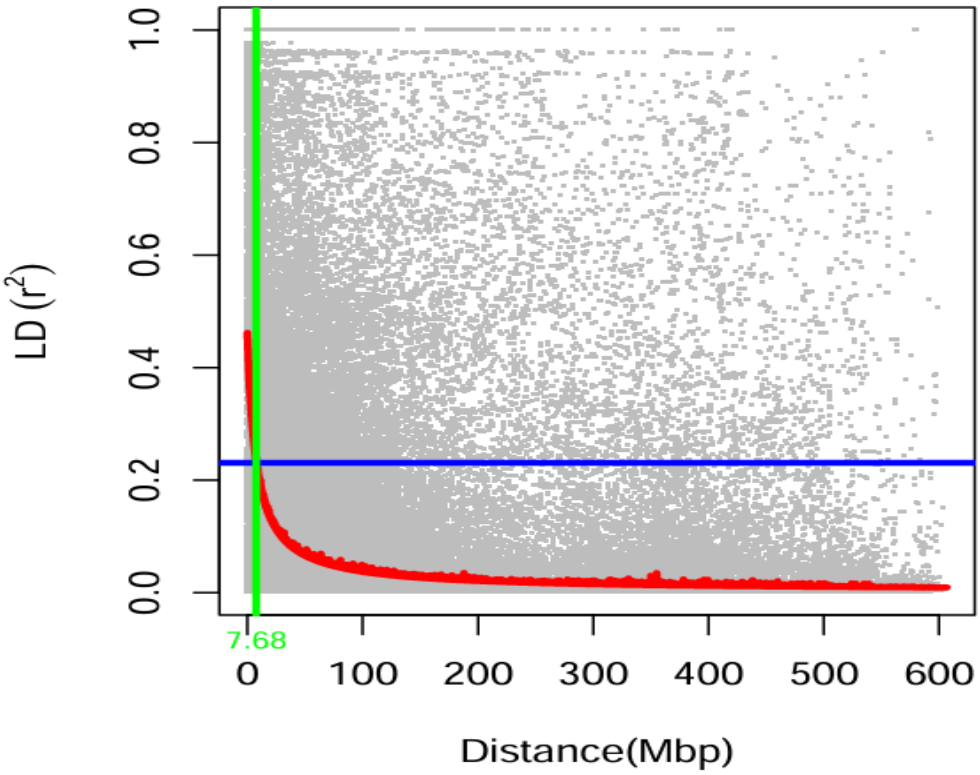


Figure 6. 2 The plot of decay in linkage disequilibrium

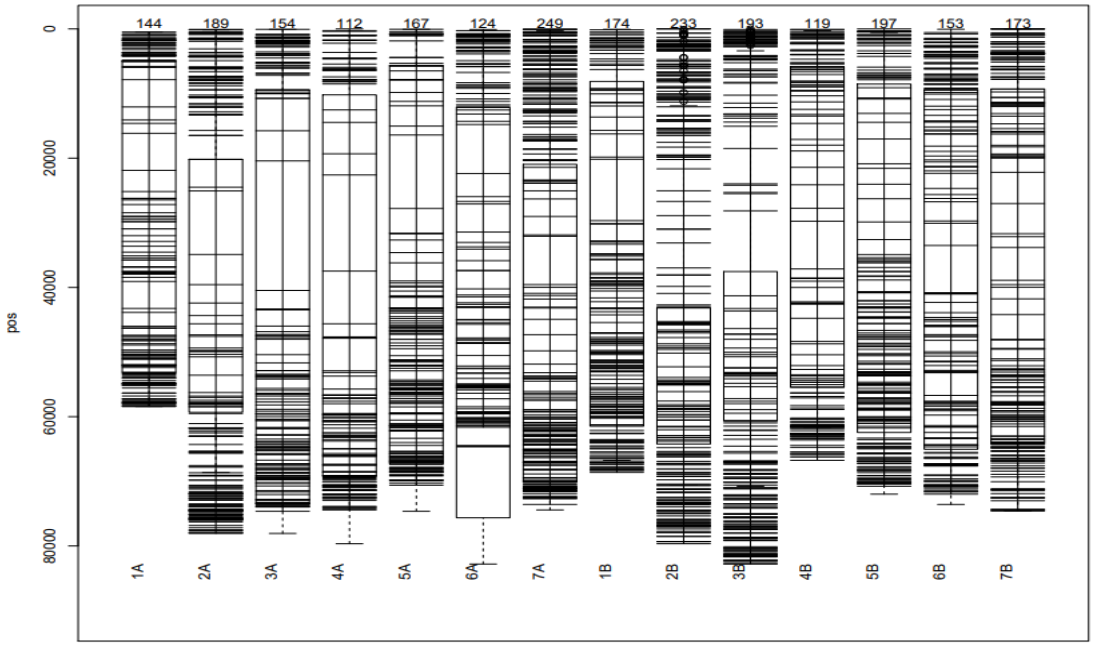


Figure 6. 3 The distribution of significant marker pairs across the genomes of tetraploid wheat

6.3.4. Marker Traits Association (MTA)

The reported SNPs included the marker-trait associations (MTAs) that surpassed a Bonferroni-correction significance threshold of 0.05. The GWAS identified 44 SNPs that were significantly (FDR <0.05) associated with some of the traits under study including the number of seeds per spikelets (SdPSp), plant height (PH), days to booting (DB), days to maturity (DM), thousand kernel weight (TKW), zeleny index (ZI), vitreousness percentage (VITP), grain yield (GY), harvest index (HI), white seed color, purple seed color, red seed color, white glume color, short awns, conspicuous awns, very lax spike density, and very dense spike density (Table 6.3). The BLUE did not provide MTAs at the stringent Bonferroni significance threshold for biomass yield, the number of spikelets per spike, the number of effective tillers per plant, spike length, gluten content, and protein content. The putative QTLs were identified across chromosome 1A, 2A, 2B, 3A, 3B, 4B, 5A, 5B, 6A, 6B, and 7B.

Table 6. 3 GWAS result obtained using GAPIT for phenotypic and grain quality traits showing significant marker trait association

SNP	Chr	Pos	P.value	FDRadj	MAF	traits	SNP	Chr	Pos	P.value	FDRadj	MAF	traits
561	2B	91084823	3.10E-09	6.2E-08	0.07609	SdPSp	971	3B	5.33E+08	2.65E-06	2.38E-05	0.47826	VITP
1595	5B	5.50E+08	5.40E-06	5.4E-05	0.4375	SdPSp	1477	5A	6.89E+08	1.47E-06	1.98E-05	0.03533	VITP
1948	6B	7.05E+08	5.46E-06	3.64E-05	0.4837	SdPSp	1565	5B	4.91E+08	1.49E-07	4.02E-06	0.07337	VITP
357	2A	75591116	8.14E-07	5.09E-06	0.47826	PH	907	3B	13884562	3.36E-08	6.22E-07	0.06522	GY
478	2A	7.47E+08	1.76E-11	4.39E-10	0.04891	PH	1916	6B	5.91E+08	4.36E-09	1.61E-07	0.44565	GY
659	2B	6.33E+08	1.76E-06	8.8E-06	0.11141	PH	70	1A	3.69E+08	1.66E-09	5.14E-08	0.46739	HI
1255	4B	3.90E+08	1.84E-05	6.56E-05	0.0625	PH	1892	6B	4.73E+08	3.92E-07	6.07E-06	0.47554	HI
1601	5B	5.63E+08	2.60E-09	3.26E-08	0.38587	PH	1017	3B	7.10E+08	6.80E-08	1.49E-06	0.16848	WSDC
1680	5B	7.04E+08	3.45E-09	2.87E-08	0.41848	PH	856	3A	6.73E+08	1.59E-05	0.000119	0.29348	PSDC
2230	7B	37247484	4.80E-06	2E-05	0.25815	PH	2339	7B	6.40E+08	2.74E-08	4.12E-07	0.1875	PSDC
73	1A	3.79E+08	6.61E-06	3.3E-05	0.10598	DB	2152	7A	6.57E+08	5.27E-07	6.85E-06	0.3587	RSDC
597	2B	2.02E+08	6.53E-06	3.81E-05	0.07337	DB	582	2B	1.54E+08	1.75E-07	4.89E-06	0.44294	GCW
1327	5A	18572611	6.44E-06	4.51E-05	0.33152	DB	1530	5B	3.57E+08	2.17E-06	3.04E-05	0.17663	GCW
1679	5B	7.03E+08	1.95E-06	1.7E-05	0.46739	DB	1261	4B	5.04E+08	3.51E-07	7.02E-06	0.10054	AwndSA
1820	6B	28685444	1.44E-06	1.68E-05	0.08424	DB	1261	4B	5.04E+08	3.51E-07	7.02E-06	0.10054	AwndCA
1856	6B	1.42E+08	1.26E-07	4.41E-06	0.45924	DB	74	1A	3.84E+08	3.92E-19	9.8E-18	0.10326	SpDenVL
1916	6B	5.91E+08	4.90E-07	8.58E-06	0.44565	DB	105	1A	5.21E+08	3.09E-09	1.93E-08	0.13859	SpDenVL
1264	4B	5.35E+08	3.05E-07	1.52E-05	0.05163	DM	498	2A	7.63E+08	1.63E-05	5.82E-05	0.07337	SpDenVL
1843	6B	80781189	7.83E-06	0.000196	0.05163	DM	851	3A	6.57E+08	6.21E-12	7.77E-11	0.39946	SpDenVL
340	2A	42932966	5.23E-06	2.76E-08	0.43207	TKW	1449	5A	6.58E+08	3.17E-11	2.64E-10	0.47283	SpDenVL
561	2B	91084823	1.73E-09	5.45E-06	0.07609	TKW	1757	6A	4.85E+08	1.70E-07	8.52E-07	0.11957	SpDenVL
805	3A	4.80E+08	6.81E-07	2.79E-05	0.24728	TKW	1764	6A	5.33E+08	2.81E-06	1.17E-05	0.46467	SpDenVL
857	3A	6.74E+08	2.32E-07	1.07E-05	0.43478	ZI	607	2B	3.81E+08	2.29E-07	5.72E-06	0.36685	SpDenVD

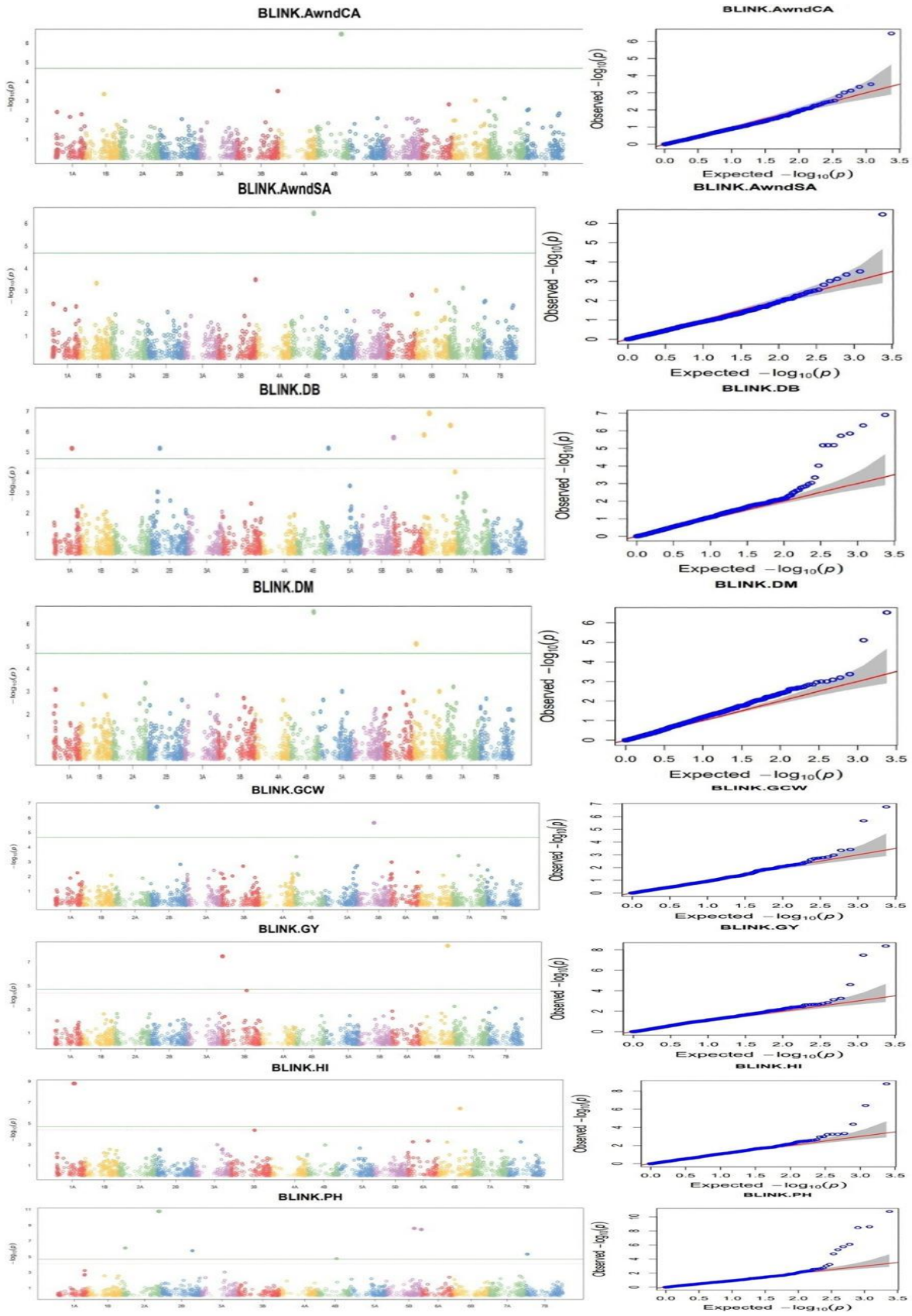
Remarks: SdPSp= number of seeds per spikelet, PH=Plant height, DB=days to booting, DM=days to maturity, TKW= thousand kernel weight, ZI = zeleny index, GY= grain yield, HI= harvest index, VITP= vitreousness percentage, WSDC= white seed color, PSDC=purple seed color, RSDC= red seed color, GCW=white glume color, AwndSA= conspicuous awn, AwndCA=short awns, SpDenVL= very lux spike density, and SpDenVD= very dense spike density.

The Genome wide association (GWA) summary is given Annex 4 and Manhattan plots and Q-Q plots of the significant markers are given Figure 6.4. Genomic positions of detected QTLs were given in Figure 6.5. The GWA scan for grain yield identified two MTAs on chromosomes 3B and 6B respectively at 13.88 Mbp and 590.91 Mbp with a respective positive (0.53) and negative (-0.80) effect. The total phenotypic variance explained by SNPs for grain yield ranged from 19.91 % for the allele 3064516|F|0-19: T>A-19: T>A on chromosome 3B at 13.88 Mbp to 43.72 % for the allele 4004985|F|0-18: A>T-18: A>T on chromosome 6B at 590.91 Mbp. Similar to this study, earlier studies reported a QTL for grain yield on chromosomes 3B (Sukumaran *et al*., 2014; Lopes *et al*., 2014) respectively at 71 cM and 278.7 cM in spring wheat, in bread wheat (Neumann *et al*., 2011) at 66.1 cM, in bread wheat at 285.77 Mbp and 697.6 Mbp (Jamil *et al*., 2019), on chromosome 3B (at 8cM) in durum wheat (Maccaferri *et al*., 2008), on chromosome 3B in durum wheat (Roncallo *et al*., 2017), on chromosome 6B at 90.97 cM (Ain *et al*., 2015), and on chromosome 6B (664,500, 180-664, 501,715 bp) in bread wheat (Rahimi *et al*., 2019). Moreover, Neumann *et al*. (2011) reported QTLs associated with grain yield along with other traits on chromosome 6B at 70.6 cM.

The GWA scan resulted in three MTAs (FDR < 0.05) for the number of seeds per spikelet on chromosomes 2B, 5B, and 6B respectively at position 91.08 Mbp, 550.32 Mbp, and 705.49 Mbp with a respective negative (-0.33), positive (0.28), and negative (-0.43) effects. The total phenotypic variance explained by SNPs for the number of seeds per spikelets ranged from 9.16 % for the allele 3947617|F|0-6:T>G-6:T>G on chromosome 2B at 91.08 Mbp to 48.33 % for the allele 1109894|F|0-17:C>G-17:C>G on chromosome 6B at 705.49 Mbp. In line with this study, QTLs associated with the number of seeds per spikelet were reported on chromosome 5B (637,387, 009-637, 389,605bp) (Rahimi *et al*. (2019) and on chromosome 2B (Roncallo *et al*., 2017); however, no report indicate QTL associated with the number of seeds per spikelet on chromosome 6B indicating this QTL could be a novel.

Plant height resulted in seven MTAs on chromosomes 2A at 75.59 Mbp and 746.61 Mbp with negative effects (-5.83 and -7.35), 2B at 632.60 Mbp with a negative effect (-0.77), 4B at 390.40 Mbp with a negative effect (-4.86), 5B at 562.79 Mbp and 703.89 Mbp with a respective negative and positive effect (-3.25 and 3.31), and 7B at 73.25 Mbp with a positive effect (2.31). The total phenotypic variance explained by SNPs for plant height ranged from 1.70 % for the allele 3021533|F|0-36:C>T-36:C>T on chromosome 7B at

37.247 Mbp to 19.45 % for the allele 1127704|F|0-29: A>C-29: A>C on chromosome 2A at 746.61 Mbp. In line with this study, QTLs associated with plant height were reported on chromosome 2A (191.58 cM, 192.76 cM), 2B (78.69 cM, 167.12 cM), and 5B (102.5 cM, 103.24 cM) in wheat cultivars from Pakistan (Ain *et al*., 2015), on chromosome 2B (142 cM) (Assanga *et al*., 2017), on chromosome 4B (95.45 Mbp) (Jamil *et al*., 2019), on chromosome 7B in wheat under late-sown (Kumar *et al*., 2020), on chromosome 2B, 5B and 7B in wheat under non-stressed conditions (Mwadzingeni *et al*., 2017), on chromosome 5B (513,646, 921-513, 649,139bp) (Rahimi *et al*., 2019), on chromosome 2B (149 cM) in durum wheat (Maccaferri *et al*., 2008), on chromosome 2B and 7B in tetraploid wheat (Tzarfati *et al*., 2014), on chromosome 2A and 4B in durum wheat (Mengistu *et al*., 2016), and on chromosome 4B in durum wheat (Wang *et al*., 2019).


BLINK.AwndCA
BLINK.AwndSA
BLINK.DB
BLINK.DM
BLINK.GCW
BLINK.GY
BLINK.HI
BLINK.PH
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

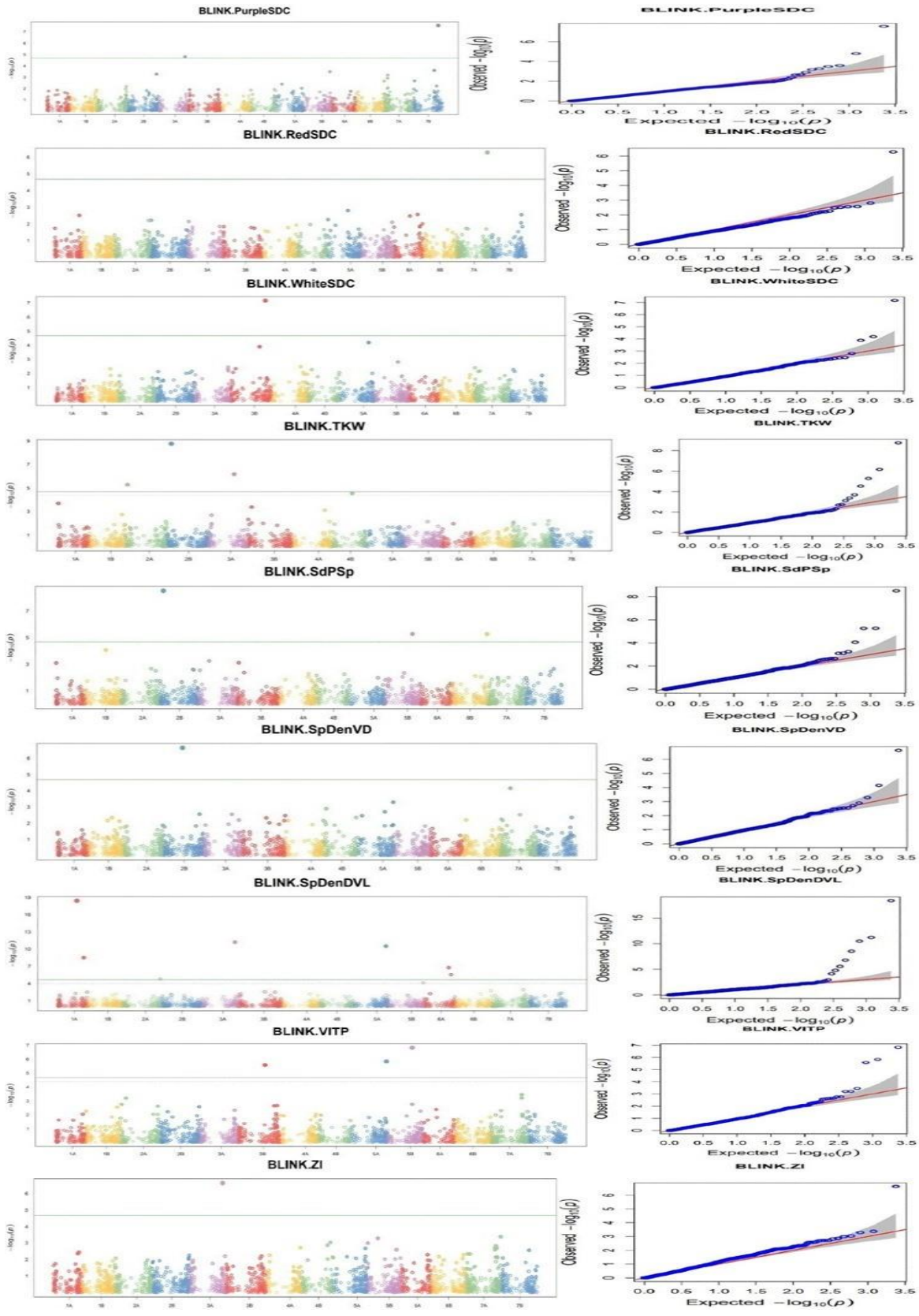


Figure 6. 4 Manhattan plots and Q-Q plots for the GWA scan on phenotypic traits and grain quality traits. Each dot represents a SNP. On the X-axis is the genomic position of the SNPs on the corresponding chromosomes indicated in different colours.Y-axis is the -log10 of the P-value depicting the significance of the association test. The horizontal green

line is the Bonferroni correction significance threshold used in the association studies of phenotypic traits and grain quality traits. The quantile-quantile plots indicate how well the GWAS model accounted for population structure and kinship for each of the traits. In each plot, the observed –log (P values) from the fitted GWAS models (y-axis) are compared with their expected value (x-axis) under the null hypothesis of no association with the trait. Each blue dot represents a single nucleotide polymorphism and the red line is the model for no assosaation. Remarks: ZI = Zeleny Index, VITP = vitreousness percentage, GY = grain yield (t/ha), TKW = thousand kernel weight, HI = harvest index, DB = days to booking, DM = days to maturity, PH = plant height, SpDenVD = very dense spike density, SpDen VL = very lux spike density, GCW = white glume color, SdPSp = number of seeds per spikelet, SDC = seed color, CA = conspicuous awns, SA = short awns, GCW = white glume color,and BLINK = Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway.

The GWA scan for days to booting identified seven MTAs three of which are on chromosomes 6B at positions 28.69 Mbp, 142.46 Mbp, and 590.91 Mbp with a respective positive (2.36), negative (-3.11), and positive (2.87) effect, one on chromosome 5B at unidentified position, and the remaining three are on chromosomes 1A, 2B, and 5A at positions 379.46 Mbp, 202.16 Mbp, and 18.57 Mbp with a respective negative (-2.57), positive (1.24), and positive (1.08) effect. The total phenotypic variance explained by SNPs for days to booting ranged from 2.08 % for the allele 978646|F|0-67: C>A-67: C>A on chromosome 1A at 379.46 Mbp to 21.22 for the allele 4004985|F|0-18: A>T-18: A>T on chromosome 6B at 590.91 Mbp. In line with the present study, study by Mengistu *et al*. (2016) detected MTA associated with days to booting on chromosome 2B (80.6 cM) in Ethiopian durum wheat and Mehrabi *et al*., (2020) identified MTA associated with days to booting on chromosome A, 2B, 5B, 5A, and 6B. Furthermore, the annotation of 5B did not result in any genes associated with days to booting in the QTL region and this might be an indication that this QTL is a novel. The GWA scan for days to maturity, on the other hand, identified two MTAs on chromosomes 4B and 6B at 535.09 Mbp and 80.78 Mbp with a respective positive (1.95) and negative effect (-1.49). The phenotypic variance explained by SNPs for days to maturity ranges from 18.07% at the allele 2260087|F|0-24: A>G-24: A>G on chromosome 6B at 80.71 Mbp to 22.42% at the allele 1216338|F|0-15: C>T-15: C>T on chromosome 4B at 535.089 Mbp. The GWA scan for days to maturity identified two MTAs on chromosomes 4B and 6B. In line with this study, QTLs associated with days to maturity were detected on chromosomes 4B (Sukumaran *et al*., 2014) and 6B (Mwadzingeni *et al*., 2017).

The GWA scan for harvest index identified two MTAs on chromosome 1A and 6B at 368.867 Mbp and 473.085 Mbp with respective negative (-8.44) and positive (4.31) effect.

The total phenotypic variance explained by SNPs for harvest index ranged from 11.03 % for the allele 1093499|F|0-23: G>C-23: G>C on chromosome 1A at 368.867 Mbp to 33.93 % for the allele 1096095|F|0-56: G>A-56: G>A on chromosome 6B at 473.085 Mbp. In line with this study, Neumann *et al*. (2011) reported QTL associated with harvest index on chromosome 1A (41.5 cM), on chromosome 1A (201-204 cM) (Assanga *et al*., 2017) in a US popular winter wheat, and on chromosome 6B (83.84 cM) in wheat cultivars from Pakistan (Ain *et al*., 2015).

The GWA scan for thousand kernel weight identified three MTAs on chromosomes 2A, 2B, and 3A respectively at position 42.93 Mbp, 91.08 Mbp, and 479.78 Mbp with a respective positive (1.78), negative (-2.61), and positive (1.55) effect. The total phenotypic variance explained by SNPs for TKW ranged from 3.90 % for the allele 1219989|F|0-38: A>G-38: A>G on chromosome 3A at 479.775 Mbp to 30.49 % for the allele 3947617|F|0-6: T>G-6: T>G on chromosome 2B at 91.08 Mbp. In line with this study, QTLs associated with TKW were reported on chromosome 2B (165.5 cM) (Assanga *et al*., 2017) in a US popular winter wheat, on chromosome 2B (134,240, 300-134, 249,722 bp) in bread wheat under rain-fed conditions (Rahimi *et al*., 2019), on chromosome 2B (52.0 cM) in bread wheat (Neumann *et al*., 2011), on chromosomes 2A (164.87 cM) and 3A (79.43 cM) in wheat cultivars from Pakistan (Ain *et al*., 2015), on chromosome 2A (752.87 Mbp) and 2B (235.16 Mbp) in bread wheat (Jamil *et al*., 2019), and on chromosome 2A (0.0-37.0 cM) in bread wheat under normal conditions (Sara *et al*., 2020).

The GWA scan for zeleny index identified one MTA on chromosome 3A at 673.588 Mbp with a negative (-0.721) effect and the total phenotypic variance explained by SNPs for it being 37.34 % for the allele 4909545|F|0-37: C>G-37: C>G. The GWA scan for vitreous percentage identified three MTAs on chromosomes 3B, 5A, and 5B respectively at 532.749 Mbp, 688.98 Mbp, and 491.07 Mbp with respective positive effects of 10.78, 14.08, and 9.76. There is no report indicating the identified loci associated with zeleny index implying that they are novel. The total phenotypic variance explained by SNPs for vitreousness percentage ranged from 13.62 % for the allele 1096353|F|0-17: G>A-17: G>A on chromosome 5A at 688.98 Mbp to 15.58 % for the allele 3026356|F|0-16: T>C-16: T>C on chromosome 5B at 491.07. In line with this study, QTLs associated with vitreousness were reported on chromosome on 5B (Johnson *et al*., 2019) and on chromosome 5A and 5B (Sun et al. 2018); however, there is no report indicating 3B is associated with vitreousness indicating that it is novel.

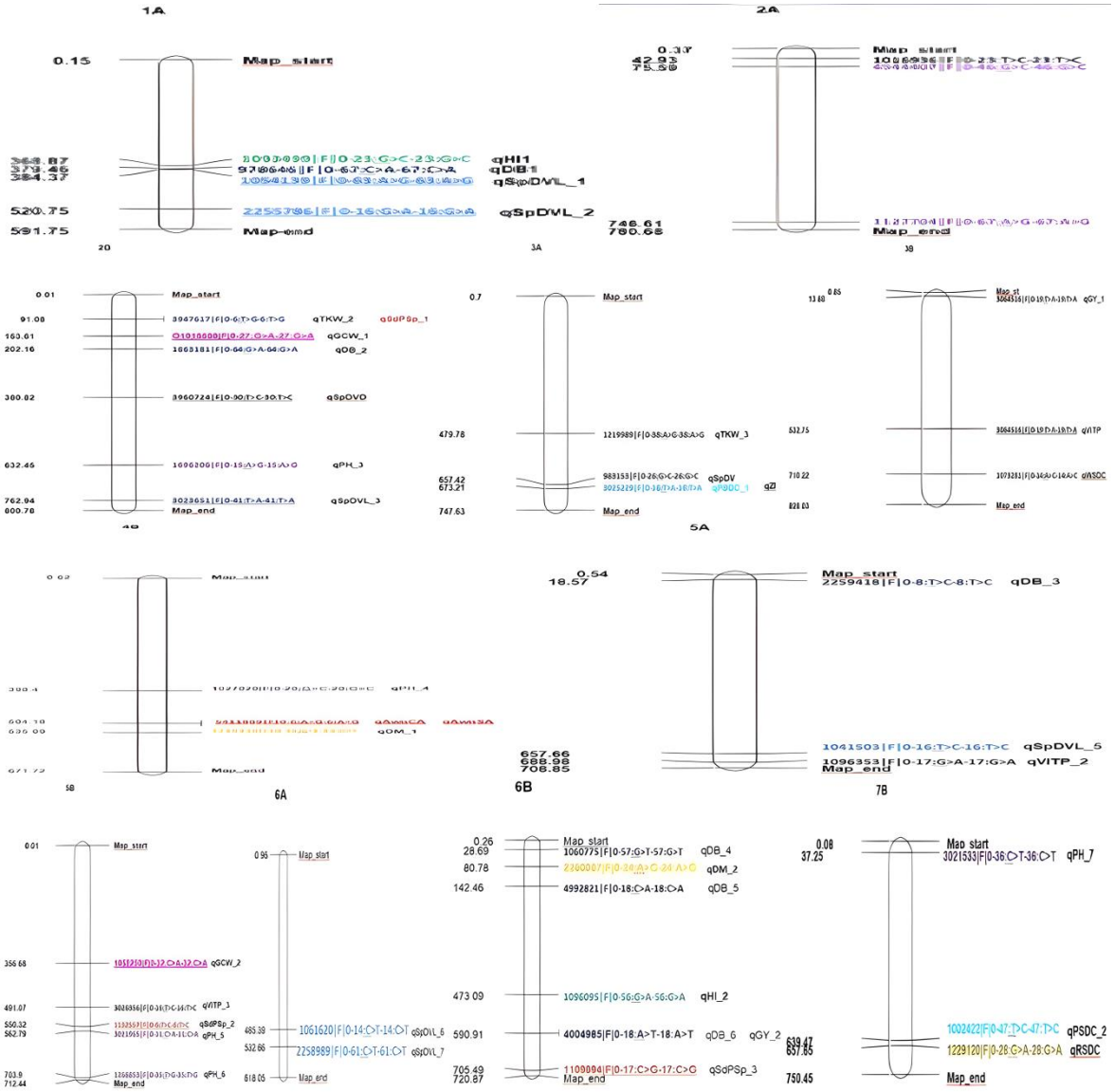


Figure 6. 5 Genomic positions of detected putative QTLs associated with phenotypic and grain quality traits. Significant DArTSeq SNPs are presented according to their physical positions on chromosomes in millions base pairs. The putative QTLs identified in this study for the MTAs are indicated on the right sides of the bars. Underlined MTAs on the right sides of the chromosomes could be potentially novel loci in this study.

The GWA scan for white seed color identified one MTA on chromosome 3B at 710.22 Mbp with a positive effect (0.18) and the total phenotypic variance explained by the SNPs for this trait was 36.14 % for the allele 1073281|F|0-14: A>C-14: A>C. The GWA scan for red seed color also identified one MTA on chromosome 7A at 656.649 Mbp with a positive effect (0.36) and the total phenotypic variance explained by the SNPs for this trait 19.72%. The GWA scan for purple seed color, on the other hand, identified two MTAs on chromosome 3A and 7B respectively at 673.21 Mbp and 639.738 Mbp with a respective negative (-0.64) and positive (0.09) effect. The total phenotypic variance explained by the

SNPs for purple seed color ranged from 6.02 % for the allele 3025229|F|0-18: T>A-18: T>A at chromosome 3A to 16.75 for the allele 1000140|F|0-25: C>G-25: C>G at chromosome 7B. In line with this study, Alemu *et al*. (2020) reported major QTLs for redness and brightness on chromosome 2A; four nominal QTLs on chromosomes 7A and 7B for brightness and five nominal QTLs for yellowness on chromosomes 3A and 7B.

The GWA scan for very lax spike density identified seven MTAs, two on chromosome 1A at 384.37 Mbp and 520.85 Mbp with a respective positive effect of 0.20 and 0.08, two on chromosome 6A at 485.39 Mbp and 532.657 Mbp with a respective positive (0.08) and negative (-0.10) effect, and the remaining three on chromosome 2B, 3A, and 5A respectively at 762.937 Mbp, 657.42 Mbp, and 657.655 Mbp with a respective negative (-0.06), positive (0.11), and positive (0.15) effect. The total phenotypic variance explained by SNPs for very lax spike density ranged from 0.19 % for the allele 2258989|F|0-61: C>T-61: C>T on chromosome 6A to 32.55 % for the allele 1041503|F|0-16: T>C-16: T>C on chromosome 5A. The GWA scan for very dense spike density identified one MTA on chromosome 2B at 380.82 Mbp with a negative effect (-3.39) and the total phenotypic variance explained by the SNPs for this trait was 36.70 %. In line with this study, QTLs associated with spike density were reported on chromosome 2A, 3A, 5A, and 6A (Liu *et al*., 2020) and on chromosome 3A in US hard winter (Halder *et al*., 2023). QTLs 1A associated with very lax spike density and 2B associated with very lax and very dense spike density were not reported and could be novel.

The GWA scan for short awns as well as conspicuous awns identified one MTA on chromosome 4B at 504.18 Mbp with a respective negative (-1.62) and positive (0.69) effect and the total phenotypic variance explained by the SNPs for the allele 5411809|F|0-6: A>G-6: A>G was 46.53 %. The GWA scan for white glume color identified two MTA on chromosome 2B and 5B respectively at 153.61 Mbp and 356.68 Mbp with a respective positive effect of 0.33 and 0.28. The total phenotypic variance explained by the SNPs ranged from 16.47 % for the allele 1058250|F|0-32: C>A-32: C>A on chromosome 5B at 356.68 Mbp to 18.10 % for the allele 1015508|F|0-27: G>A-27: G>A on chromosome 2B.

6.3.5. Pleiotropic locus: A pleiotropic locus is associated and affects the expression of more than one phenotypic trait (Mwadzingeni et al. 2017). In this study, two pleiotropic loci were identified including the marker 3947617|F|0-6: T>G-6: T>G located on chromosome 2B for seed per spikelet (-0.32733) and thousand kernel weight (-2.60604)

having a negative effect in both traits and 5411809|F|0-6:A>G-6:A>G on chromosome_4B for conspicuous (-1.61935) and short awns (0.694008).

6.3.6. Functional annotation: The functional association of the identified QTLs for agronomic and grain quality traits was further investigated by annotating genes found in the QTL regions on the recently released IWDSC RefSeq Annotations v2. The analysis resulted in several high confidence genes associated with grain yield, TKW, days to booting, days to maturity, spike density, plant height, awndness, seed color, vitreousness percentage, and zeleny index (Annex 5). These high confidence genes encode various proteins, including Expansin, glycosyltransferases, F-box family protein, Myb transcription factors, Cyclin-dependent kinase inhibitor, WRKY transcription factor, and Cytochrome P450 superfamily protein, which are involved in the regulation of various plant growth and development processes.

Expansin, which is highly involved in stem elongation in common wheat and loosens plant cell walls and involves in cell enlargement and various abiotic stresses (Yue *et al*., 2018); glycosyltransferases, members of the multi gene superfamily in plants, transfer single or multiple activated sugars to a variety of plant molecules, resulting in plant compounds glycosylation and plays a key role in maintaining cell homeostasis, thus likely participating in the regulation of plant growth, development, and in defense responses to stress environments (Wang and Hou 2016; Lim and Bowles 2004) and are involved in the biosynthesis of glycolipids, sucrose, starch, and cell-wall polysaccharides, the addition of N-linked glycans to glycoproteins or modification of plant glycoproteins, and the attachment of sugar moieties to various small molecules such as hormones and flavonoids (Keegstra and Raikhel 2001); F-box family protein, which are involved in the regulation of various developmental processes in plants, such as photomorphogenesis, circadian clock regulation, self-incompatibility, embryogenesis, seed germination, plant growth and development, and floral meristem and floral organ identity determination, responses to biotic and abiotic stress, plant secondary metabolism, hormonal responses , and senescence (Zhang *et al*., 2019; Jain *et al*., 2007); Myb transcription factors, which are involved in the regulation of plant growth and tissue/organ development such as xylem, guard cell, trichome, and root hair formation, cellular proliferation and differentiation, hormonal responses during seed development and germination, and biosynthesis of diverse secondary metabolites such as flavonoids, alkaloids, terpenoids, and phenolics, as well as abiotic and biotic stress responses and disease resistance (Li *et al*., 2022; Ambawat *et al*.,

2013; Wang *et al*., 2021); Cyclin-dependent kinase inhibitor, which is significant for cell division control and helps plant optimize its architecture (Tank and Thaker 2011), essential in orchestrating eukaryotic cell proliferation and differentiation (Clercq and Inzé 2006), play an important role in regulating cell cycle exit that precedes differentiation and in promoting differentiation in cooperation with transcription factors (Clercq and Inzé 2006), and coordinate cell division, cell expansion and organ growth with developmental and environmental cues (Kumar and Larkin 2017); WRKY transcription factor, a plant-specific transcription factor group (Li *et al*., 2020), are one of the largest families of transcriptional regulators in plants and form integral parts of signalling webs that modulate many plant processes and play roles in regulating several different plant processes such as biotic and abiotic stresses, seed development, seed dormancy and germination, and leaf senescence (Rushton *et al*., 2010); Cytochrome P450 superfamily protein, the largest enzymatic protein family in plants, which are involved in multiple metabolic pathways with distinct and complex functions, playing important roles in promoting plant growth and development and protecting plants from stresses through multiple biosynthetic and detoxification pathways (Xu *et al*., 2015), involved widely in the synthesis of secondary metabolites and regulation of plant hormone metabolism, participate in many important cell processes, play a significant role in plant defense responses (Xu *et al*.,2015; Li and Wei 2020;Gunupuru *et al*., 2018;Wang *et al*., 2023), take part in biochemical pathway such as the biosynthesis of jasmonate (JA), act in the wheat stem and/or root lignin synthesis pathway (Li and Wei 2020), are involved in diverse biological processes including biotic and abiotic stress response and herbicide tolerance (Werck-Reichhart *et al*,2000; Wang *et al*., 2020) and are positive contributors to both biotic stress resistance and grain development in wheat (Gunupuru *et al*., 2018). Overall, in the present study 44 significant MTAs, two of which are pleiotropic, were identified; out of which, thirty-five QTLs were reported while the remaining 9 QTLs are found to be novel, which need to be validated for further use in the breeding program.

6.4. Conclusion

Novel QTLs were detected in a panel of genetically rich and untapped Ethiopian tetraploid wheat germplasm in a two environment (two-location and two-year) study. Use of multi-environment, multi-year phenotypic data could reveal QTL stable across environments. The GWAS scan identified 44 significant MTAs, two of which are pleiotropic, across chromosome 1A, 2A, 2B, 3A, 3B, 4B, 5A, 5B, 6A, 6B, and 7B. Nine of these MTAs are novel while the remaining 35 were reported. We also identified candidate genes for the novel loci potentially regulating the traits. Hence, this study highlights the significance of the Ethiopian tetraploid wheat gene pool for improving tetraploid wheat globally. Furthermore, a breeding strategy focusing on accumulating favorable alleles at these loci could improve tetraploid wheat production in Ethiopia and internationally.

CHAPTER VII

7. Summary and Recommendations

There is a pressing need to increase production and productivity to feed the ever-increasing world population (FAO, 2017; FAO, 2019). Wheat is one of the commodities given priority to attain the objective of food self-sufficiency and food and nutrition security. Moreover, there is an increased requirement for high-quality wheat-based food products. Grain quality of wheat becomes central to human well-being (Kong *et al*., 2013) due to changing food habits, increasing urbanization, and trends toward raising middle-class society (Wang *et al*., 2017). Research efforts so far made contributed justifiable achievements to increase production and productivity of wheat; however, the productivity of wheat in Ethiopia is below the global and attainable yield level. Thus, continued enhancement of wheat yield is needed. In addition to enhancement of wheat yield, improving wheat grain quality is of vital value to meet the current increased demand. A continuous supply of new germplasm resources is needed to enhance wheat yield and to improve its grain quality for which the contribution of tetraploid wheat landraces is great.

Tetraploid wheat *(Triticum turgidum spp.)* species of Ethiopia hold ample genetic variation with many desirable agronomic traits, which could provide a source for improvement of wheat. Even though Ethiopia is a center of diversity for tetraploid wheat and the largest wheat producer in sub-Saharan Africa; the country remains net importer of wheat due to huge gap between production and consumption. Thus, knowing the extent of genetic diversity is vital for improvement, efficient utilization, and conservation of the germplasm. Knowledge of genetic diversity can be obtained using morphological, biochemical, and molecular markers. Furthermore, undertaking genome wide association mapping, which is a powerful technique used to identify marker-trait associations, is crucial to elucidate the genetic basis of complex quantitative traits controlling agronomic traits of economic importance. The key to modern plant breeding is elucidating the genetic basis of complex quantitative traits controlling agronomic traits of economic importance (Alemu *et al*., 2020) using recent genomic based research. This will facilitate the utilization of Ethiopian tetraploid germplasm in the crop improvement program.

The first objective of the present study was to assess the extent and pattern of diversity of tetraploid wheat germplasm of Ethiopia based on phenotypic traits. The genetic diversity study based on phenotypic traits was conducted using 4 qualitative traits and 11 quantitative traits.

In the present study, intermediate (0.60±0.16) diversity was observed in the tetraploid wheat germplasm of Ethiopia based on qualitative traits. Shannon-Weaver diversity revealed high diversity in the entire germplasm with respect to glum color (0.78), spike density (0.61), and seed color (0.86); however, low diversity was observed for awndness (0.15). All species exhibited high H' for glume color (0.77±0.01) and seed color (0.77±0.09); however, intermediate and low diversity was observed for spike density (0.57±0.02) and awndness (0.17±0.03), respectively. In the present study, the phenotypic classes observed for seed color, spike density, awndness, and glume color were respectively white, red, and purple; very lax, lax, intermediate, dense, and very dense; awnletted and conspicuous or large awns, and white, red to brown, and purple to black. Within region diversity is greater than between regions diversity implying that there are enough variations within the region as compared to between regions for seed color, glume color, awndness, and spike density.

Genetic diversity assessment based on the quantitative traits revealed highly significant variation ($p<0.001$) for all traits, except for the number of effective tillers per plant among genotypes and genotype by location interaction. This suggests that there is large possibility for genetic improvement through selection. All traits showed highly significant variation ($p<0.001$) across locations implying huge impact of test sites on traits. Phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV) values were high for grain yield, biomass yield, and harvest index, suggesting large possibility for genetic improvement through selection. The number of seeds per spikelet showed high heritability while the remaining quantitative traits showed moderate heritability, indicating the presence of alleles of breeding relevance. The number of seeds per spikelet gave a high value of genetic advance accompanied with high heritability implying high genetic gain from its selection. Thus, the top 5% of the tetraploid wheat genotypes, Alemtena, Fetan, Malefia, DZ2018, Tob66, Werer, Selam, Cocorit/71, and Ejersa, for the number seeds per spikelet could be used as potential parents for crossing to develop improved varieties for this trait. Most of the traits studied showed positive association with grain yield implying that improving either one or more of these traits could result in high grain yield.

The second aim of this study was to assess the extent and pattern of diversity based on grain quality traits. The highly significant ($P < 0.001$) variation among the genotypes for all grain quality traits studied suggests that there is a substantial genetic diversity, which is crucial for breeding programs providing a broad base for selecting superior genortypes. The highly signficant interaction of genotype by site for all grain quality traits indicates that the performance of genotypes varies across different environments highlighting the importance of considering environmental conditions in breeding decisions. Except vitreousness percentage, all the traits studied were highly significantly affected by the site of study. High range values were recorded for vitreousness percentage (69.7), zeleny index (43.71), and thousand kernel weight (21.3) implying that there is significant genetic variation within the population for these traits and that there is broader base for improvement through selection based on these traits.

The protein content showed the lowest PCV and GCV values, which is an indication of limited genetic diversity and stable expression for protein content suggesting that simple selection is ineffective for enhancing protein content. The low GCV values for protein content implies that there is less scope for selection because the genetic diversity within the present tetraploid germplasm is limited for protein content. There was high GCV value for grain yield (24.03) and intermediate GCV values for thousand kernel weight (11.58), gluten content (10.81), zeleny index (19.04), and vitreousness percentage (13.22) indicating that there is good scope for selection of these traits. Grain yield (37.18), zeleny index (23.37), and vitreousness percentage (22.82) showed high PCV implying that these traits vary significantly among genotypes, reflecting both genetic and environmental influences and that there is greater potential for selection based on these traits. The difference between PCV and GCV was low for protein content, gluten content, thousand kernel weight, zeleny index, and vitreousness percentage indicating minimum environmental influences and consequently greater role of genetic factors on the expression of these traits and intermediate for grain yield implying a high variability, which is influenced by environment for its expression.

Environmental coefficient of variation was high for grain yield (31.38) implying that environment plays significant role in influencing the variability of grain yield indicating that environmental factors such as soil quality, weather conditions, water availability, and soil fertility greatly impact grain yield. This means that differences in yield are more likely

due to external conditions rather than genetic differences among the genotypes. On the other hand, intermediate ECV was observed for vitreousness percentage (19.92), zeleny index (18.42), thousand kernel weight (13.46), and gluten content (12.20) implying that environmental factors have moderate impact on these traits indicating that moderate environmental variability can be advantageous in breeding programs for it provides a balance between stability and the potential for selecting individuals that perform well under varying conditions. The observed ECV value was low for protein content (6.73) indicating that the trait is stable across different environmental conditions. Since the trait is not significantly influenced by environmental variability, breeders can confidently select for high protein content without worrying about environmental factors distorting the results. Except grain yield, which showed high value (33.42), other traits showed lowest values for genotype by environment interaction coefficient of variation suggesting that grain yield is highly sensitive to environmental changes.

Except vitreousness percentage which had low heritability, other grain quality traits exhibited medium heritability indicating a balanced influence of genetic and environmental factors and that selective breeding can effectively enhance these traits. Heritability and genetic advance decide the success of selection in breeding programs. Vitreousness percentage had low heritability coupled with medium genetic advance while protein content had low heritability combined with low genetic advance indicating the lower success in selection for these traits.

All of the genotypes gave above 11 % protein, which is suitable for bread-making. A total of 42 genotypes, including forty landraces along with two improved varieties (Ld-357 and Coccorit/71), gave protein content higher than 14 %, which is demanded by the modern pasta manufacturing industry and is in the range (12% and 16%) of protein level for quality pasta products. On the other hand, 115 (59%) of the genotypes gave the mean value of protein content higher than 13 percent, excellent pasta cooking quality and a standard in quality throughout the grain industry. All the improved varieties along with 171 landraces gave higher than 11.50 % protein content, which is a seed protein level in the hard wheat. Four landraces (6955, 5728, 5504, and 5465) along with the improved variety Coccorit/71 gave the mean value of protein greater than 15 % and among these landraces, landraces 5728 (4.52 t/ha, 36.5 gm) and 5465 (4.45 t/ha, 37.81 gm) gave high grain yield (t/ha) and medium thousand kernel weight. Zeleny index was on average 41.57 ml with a minimum and maximum value of 19.48 ml to 63.19 ml and a high range value (43.71). Except the

landrace 5142, all the genotypes had average zeleny index above 20 ml. Vitreousness percentage was on average 68.44% with a minimum and maximum value of 28.48 % and 98.18 %, respectively, and a high value of range (69.7). The landrace 214370, which had highest vitreousness percentage had a low yield (2.21 t/ha) and was not included in the selected genotypes based on their performance. A total of 51 genotypes, including sixteen improved varieties along with 35 landraces, recorded more than 80 % vitreous percentage, which is the acceptable minimum value, and these landraces are especially suited to the production of pasta because of their highly vitreous grain or high milling yield of semolina and the remaining genotypes recorded vitreousness percentage less than 80 %, which is undesirable for pasta milling; therefore, it is recommended to incorporate these landraces in a breeding program for further selection for future release for pasta production. Vitreousness in wheat is regarded as high (above 75 %), medium (between 60 % and 75 %), and low (below 60 %) (Korkut, Lg, and Lam 2007). In the present study, 77, 63, and 56 of the genotypes had vitreousness value above 75 %, between 60 % and 75 %, and below 60 %. A total of twenty two genotypes, including eleven improved varieties (Tob66, Foka, Yerer, Quamy, Mangudo, Ginchi, Asasa, Boahi, Bekelcha, Selam, and Werer) and eleven landraces (7060, 6914, ETDW/15DZ#044, 5510, 7209, 7666, 226694, 5893, DZW17002,6975, and DZW17007), were top performing based on their grain yield and grain quality traits performance; thus, it is recommended to incorporate the landraces in a breeding program for further release and the released varieties could be utilized for wider production after evaluating their yield stability across locations. Except landraces 7060 and 226694, which had red seed color, other selected landraces and improved varieties had either white or purple (yellow) seed color, which is most preferred for pasta production. On the other hand, four landraces (7060, 7209, 6914, and ETDW/15DZ#044) along with four improved varieties (Selam, Quamy, Tob66, and Werer) could be included in a breeding program as a potential genotypes in areas where feed source is scarce due to their high average biomass yield in addition to their grain quality traits.

Principal component analysis based on phenotypic traits revealed that the first five principal components, with Eigen values greater than one combined, explained 61.21% of the total variation and 89.17 % of the total variation was contributed from the first ten components. The variation in the first principal component was mainly contributed by days to booting, the number of seeds per spikelet, days to maturity, spike density and harvest index. On the other hand, PCA based on grain quality traits revealed that the first three principal components with Eigen values greater than one combined explained 69.11% of

the total variation. The variation in the first principal component was mainly contributed by zeleny index (ml), vitreousness percentage and gluten content.

Based on regions of collections, the first two components alone explained 90.7% of the total variation based on phenotypic traits and the first three components explained 89.99% of the total variation based on grain quality traits, which suggests that a large portion of the phenotypic traits and grain quality traits variability can be captured by these two and three principal components, respectively. The results from the principal component analysis (PCA) based on altitudinal classes, where the first four components explained 99.43% of the total variation for phenotypic traits and the first three components accounted for 95.05% of the variation in grain quality traits imply that these components effectively captured the variability present in the phenotypic traits and in the grain quality traits across different altitudinal classes. The species-based PCA, on the other hand, showed that the first two principal components captured a substantial proportion of the total variation for both phenotypic and grain quality traits; thus, demonstrating a powerful simplification and summarization of trait variability, facilitating more efficient and targeted breeding strategies. By focusing on the principal components that explain the vast majority of variation, breeders can improve selection accuracy and enhance trait development.

Cluster analysis based on phenotypic and grain quality traits showed that there is no distinct regional and altitudinal grouping, which could be due to gene flow or the same selection criteria used by the farmers across regions resulting in similar phenotypes. Cluster analysis indicated that landraces from the same location do not fall in the same cluster and at the same time landraces from different localities may fall in the same cluster implying that there is no relation between geographic and genetic distances. This, in turn, implies that accessions collected from the same geographic area were not necessarily closely related, and different regions did not necessarily have different genetic backgrounds. Clustering based on phenotypic traits and grain quality traits showed different result; likewise, clustering output based on molecular analysis is also different from the result from phenotypic traits and grain quality traits. Thus, it is reliable to consider the result obtained using molecular analysis. The whole tetraploid wheat study panel was clustered into two distinct populations (landraces and released varieties) based on molecular diversity analysis underscoring the importance of leveraging both traditional genetic diversity and modern breeding advancements. This dual approach can enhance the

adaptability, resilience, and overall performance of wheat varieties. Conservation efforts should focus maintaining the genetic diversity in Pop2 is maintained and utilized effectively, for genotypes in within Pop2 is crucial for future breeding programs; however, efforts to conserve Pop1 should focus on preserving the unique alleles present in the landraces, even if the overall diversity is lower as these landraces might contain valuable traits that are not present in the improved cultivars.

The present study also attempted to assess the genetic diversity and population structure of tetraploid wheat germplasm of Ethiopia based on molecular markers. Analysis of molecular variance (AMOVA) revealed significant ($P<0.001$) variation between populations and within populations; the variation within population being higher than between populations. Fixation index (Fst) was reported to be little ($Fst < 0.05$) for the landraces (= 0.028) and moderate (Fst= 0.05 to 0.15) for the study panel (Fst = 0.117) implying no significant differentiation among populations; however, great (Fst=0.15 to 0.25) genetic differentiation was recorded for the regions of collections (Fst=0.162) and at the species (Fst=0.162) level. The higher gene flow observed in the study populations resulted in the loss of genetic differentiation among populations implying a low level of diversity among subpopulations. This result is also further verified by the cluster analysis results, based on quantitative traits and grain quality traits, which showed that there is no distinct regional and altitudinal grouping of tetraploid wheat, which could be due to gene flow, or the same selection criteria used by the farmers across regions resulting in similar phenotypes.

The genetic indices of the study panel revealed that pop2, which comprised 19 improved cultivars and 12 landraces, was more diverse than pop1, which comprised of 98.04% landraces along with 1.96% released varieties. The PCA results based on molecular markers highlight the distinct separation between improved cultivars and landraces, emphasizing the potential for crossbreeding to enhance genetic diversity and trait performance. The diversity within landraces, despite the lack of clear clustering, offers valuable genetic resources for breeding programs. The clear separation between released varieties and landraces implies that breeding strategies can be tailored to combine the best traits from both groups, leveraging the improved traits of released varieties and the diverse genetic backgrounds of landraces. The lack of clear groupings within landraces indicates a rich genetic diversity that can be tapped into for developing new varieties with unique

traits. The first two components together explained 77 % (PC1 = 65 %, PC2 = 12 %) and 68 % (PC1 = 48 %, PC2=20 %) of the total variation in the tetraploid wheat study panel and landraces, respectively indicating that the major sources of variability can be captured by these two components. DArTSeq markers revealed maximum (0.387) genetic distance between ETW115 (7683) and ETW135 ('MCD11-20, MCD8-2, ETDW/15DZ019,……') in the study panel. This is an indication of significant genetic differentiation between these two genotypes; and, such a large distance suggests substantial genetic diversity, making these genotypes valuable for breeding programs aiming to introduce variability. Similarly, Tesfaye and Bekelcha, had maximum (0.553) genetic distance, which shows a high level of genetic differentiation; there, this significant diversity can be exploited to combine desirable traits from both varieties.

The total number of different alleles reported both in pop1 (4762, 4624) and pop2 (4708, 4760) in the present study is an indication of rich genetic diversity in both the study panel and the landraces. The presence of private alleles in 127 genotypes highlights the potential for identifying and utilizing unique genetic variations that could contribute to key adaptive traits. These alleles can be targeted in breeding programs to introduce desirable traits into new varieties and hence to develop varieties with improved resilience and performance. So, the rich genetic diversity and the private alleles reported in the present study offer significant opportunities for breeding programs to enhance crop performance and resilience. Therefore, strategically exploiting these genetic resources in a breeding program can enhance the development of improved varieties that are better suited to meet future agricultural challenges.

The fourth objective of the present study was to detect the association of phenotypic traits and grain quality traits with DArTSeq markers. The GWAS scan identified 44 significant MTAs across chromosomes; two of which are pleiotropic. Nine of these MTAs are novel while the remaining thirty-five were reported. We also identified candidate genes for the novel loci potentially regulating the traits. Hence, this study highlights the significance of the Ethiopian tetraploid wheat gene pool for improving tetraploid wheat globally.

In conclusion, the present study revealed high genetic diversity in Ethiopian tetraploid wheat (*Triticum turgidum spp.*) germplasm, which could be utilized for future wheat breeding program. The significant variation among genotypes and substantial GxE interaction observed in the present study in both the phenotypic traits and grain quality traits underscores the necessity of comprehensive breeding strategies that account for both

genetic potential and environmental influences. This dual approach can lead to the development of robust and high-quality grain varieties that perform well across diverse conditions. Conservation strategies such as preserving isolated subpopulations or reducing gene flow must be devised to maintain or enhance genetic diversity owing to the higher gene observed in the present study. Moreover, the higher genetic variation observed within populations than between populations reveal that more attention should be given to individual accessions within populations to explore the existing genetic diversity. It is recommended to incorporate the landraces (ETW115 (7683) and ETW135 ('MCD11-20, MCD8-2, ETDW/15DZ019,……') and the released varieties (Tesfaye and Bekelecha) which had maximum genetic distance to combine desirable traits from these genotypes.

The present study revealed the existence of rich genetic diversity and 127 genotypes with one or more private alleles at 755 loci. Moreover, the GWAS scan identified nine MTAs, which are novel. Hence, this study highlights the significance of the Ethiopian tetraploid wheat gene pool for improving tetraploid wheat globally. Furthermore, a breeding strategy focusing on accumulating favorable alleles at these loci could improve tetraploid wheat production in Ethiopia and internationally. Furthermore, the diversity and effectiveness of the newly identified QTLs need to be functionally validated to guarantee their use for marker-assisted selection.

8. References

Abate, Dubale. 2018. Review on Market Chain Analysis of Wheat in Ethiopia. *Industrial Engineering Letters* 2 (6): 94–105. https://doi.org/10.7176/iel/9-1-02.

Abdurakhmonov, I.Y, and Abdusattor Abdukarimov. 2008. Application of Association Mapping to Understanding the Genetic Diversity of Plant Germplasm Resources. *International Journal of Plant Genomics* 2008. https://doi.org/10.1155/2008/574927.

Abebe, Admas Alemu, and Tesfaye Molla Desta. 2017. Genetic Variability and Character Association of Some Local Wheat Varieties (*Triticum species*) Using Agromorphological Traits Grown in South Gondar Zone , Ethiopia. *African Journal of Biotechnology* 16 (44): 2100–2105. https://doi.org/10.5897/AJB2015.15018.

AGTEC-Org. 2011. Agronomical and Technological Methods to Improve ORganic Wheat Quality.

Ahuja, M.R., and Shri Mohan Jain. 2015. *Genetic Diversity and Erosion in Plants: Indicators and Prevention*. Vol. 1. https://doi.org/10.1007/978-3-319-25637-5.

Ain, Qurat Ul, Awais Rasheed, Alia Anwar, Tariq Mahmood, Muhammad Imtiaz, Tariq Mahmood, Xianchun Xia, Zhonghu He, and Umar M. Quraishi. 2015. Genome-Wide Association for Grain Yield under Rainfed Conditions in Historical Wheat Cultivars from Pakistan. *Frontiers in Plant Science* 6 (Article 743): 1–15. https://doi.org/10.3389/fpls.2015.00743.

Akbari, Mona, Peter Wenzl, Vanessa Caig, Jason Carling, Ling Xia, Shiying Yang, Grzegorz Uszynski, et al. 2006. Diversity Arrays Technology (DArT) for High-Throughput Profiling of the Hexaploid Wheat Genome. *Theoretical and Applied Genetics* 113 (8): 1409–20. https://doi.org/10.1007/s00122-006-0365-4.

Alamerew, Sentayehu, Sabina Chebotar, Xiuqiang Huang, and Marion Roder. 2004. Genetic Diversity in Ethiopian Hexaploid and Tetraploid Wheat Germplasm Assessed by Microsatellite Markers. *Genetic Resources and Crop Evolution* 51 (5): 559–67.

Alemu, Admas, Tileye Feyissa, Tesfaye Letta, and Bekele Abeyo. 2020. Genetic Diversity and Population Structure Analysis Based on the High Density SNP Markers in Ethiopian Durum Wheat (*Triticum turgidum* Ssp. Durum). *BMC Genetics* 21 (18): 1–12. https://doi.org/10.1186/s12863-020-0825-x.

Alemu, Admas, Tileye Feyissa, Marco Maccaferri, Giuseppe Sciara, Roberto Tuberosa, Karim Ammar, Ayele Badebo, Maricelis Acevedo, Tesfaye Letta, and Bekele Abeyo. 2021. Genome-Wide Association Analysis Unveils Novel QTLs for Seminal Root System Architecture Traits in Ethiopian Durum Wheat. *BMC Genomics* 22 (1): 1–16. https://doi.org/10.1186/s12864-020-07320-4.

Alemu, Admas, Tileye Feyissa, Roberto Tuberosa, Marco Maccaferri, Giuseppe Sciara, Tesfaye Letta, and Bekele Abeyo. 2020. Genome-Wide Association Mapping for Grain Shape and Color Traits in Ethiopian Durum Wheat (*Triticum turgidum* Ssp. Durum). *Crop Journal* 8: 757–68. https://doi.org/10.1016/j.cj.2020.01.001.

Alemu, Admas, Sufian Suliman, Adel Hagras, Sherif Thabet, Ayed Al-abdallat Awadalla, Abdalla Abdelmula, and Wuletaw Tadesse. 2021. Multi-Model Genome-Wide Association and Genomic Prediction Analysis of 16 Agronomic , Physiological and Quality Related Traits in ICARDA Spring Wheat. *Euphytica* 217 (205): 1–22. https://doi.org/10.1007/s10681-021-02933-6.

Alemu, Sisay Kidane, Ayele Badebo Huluka, Kassahun Tesfayei, Teklehaimanot Haileselassie, and Cristobal Uauy. 2021. Genome-Wide Association Mapping Identifies Yellow Rust Resistance Loci in Ethiopian Durum Wheat Germplasm. *PLoS ONE* 16 (5 May 2021): 1–28. https://doi.org/10.1371/journal.pone.0243675.

Alemu, Tolesa. 2020. Wheat Production Efficiency and Profitability Wheat Production. Addis Abeba.

Alemu, Tolesa, Habtemariam Zegeye, Daniel Kassa, Dawit Asnake, Tafesse Solomon, and Alemayo Asefa. 2019. Wheat Product Concepts Validation and Assessment of Dissemination and Utilization Constraints. Addis Abeba.

Allard, Robert W. 1960. *Principles of Plant Breeding*. John Wiley and Sons inc.

Ambawat, Supriya, Poonam Sharma, Neelam R. Yadav, and Ram C. Yadav. 2013. MYB Transcription Factor Genes as Regulators for Plant Responses: An Overview. *Physiology and Molecular Biology of Plants* 19 (3): 307–21. https://doi.org/10.1007/s12298-013-0179-1.

Amiri, Reza, Shahryar Sasani, Saeid Jalali-Honarmand, Ali Rasaei, Behnaz Seifolahpour, and Sohbat Bahraminejad. 2018. Genetic Diversity of Bread Wheat Genotypes in Iran for Some Nutritional Value and Baking Quality Traits. *Physiology and Molecular Biology of Plants* 24 (1): 147–57. https://doi.org/10.1007/s12298-017-0481-4.

Aoun, Meriem, Matthew Breiland, M Kathryn Turner, Alexander Loladze, Shiaoman Chao, Steven S Xu, Karim Ammar, James A Anderson, and James A Kolmer. 2016. Genome-Wide Association Mapping of Leaf Rust Response in a Durum Wheat Worldwide Germplasm Collection. *Plant Genome* 9 (NO.3): 1–24. https://doi.org/10.3835/plantgenome2016.01.0008.

Araus, José Luis, Gustavo A Slafer, Conxita Royo, and M Dolores Serret. 2008. Critical Reviews in Plant Sciences Breeding for Yield Potential and Stress Adaptation in Cereals Breeding for Yield Potential and Stress Adaptation in Cereals. *Plant Sciences* 37: 377–412. https://doi.org/10.1080/07352680802467736.

Arega, G, M Hussein, and H Singh. 2010. Genotypic Variability, Heritability, Genetic Advance and Associations among Characters in Ethiopian Durum Wheat (*Triticum durum* Desf.) Accessions." *East African Journal of Sciences* 4 (1): 27–33. https://doi.org/10.4314/eajsci.v4i1.71521.

Asmamaw, Meseret, Gemechu Keneni, and Kassahun Tesfaye. 2019. Genetic Diversity of Ethiopian Durum Wheat (*Triticum durum* Desf .) Landrace Collections as Reveled by SSR Markers. *Advances in Crop Science and Technology* 7 (1): 1–9. https://doi.org/10.4172/2329-8863.1000413.

———. 2020. Genetic Diversity of Ethiopian Durum Wheat (*Triticum durum* Desf) Landrace Collections as Reveled by Morphological Markers. *Journal of Plant Breeding and Crop Science* 12 (04): 258–68. https://doi.org/10.4172/2329-8863.1000413.

Assanga, Silvano O., Maria Fuentealba, Guorong Zhang, Chor Tee Tan, Smit Dhakal, Jackie C. Rudd, Amir M.H. Ibrahim, et al. 2017. Mapping of Quantitative Trait Loci for Grain Yield and Its Components in a US Popular Winter Wheat TAM 111 Using 90K SNPs. *PLoS ONE* 12 (12): 1–21. https://doi.org/10.1371/journal.pone.0189669.

Ayana, Amsalu, and Endashaw Bekele. 1999. Multivariate Analysis of Morphological Variation in Sorghum (*Sorghum bicolor* (L.) Moench) Germplasm from Ethiopia and Eritrea Amsalu. *Genetic Resources and Crop Evolution* 46: 273–84. https://doi.org/10.1023/A:1003998313302.

———. 2000. Geographical Patterns Ofmorphological Variation in Sorghum (*Sorghum bicolor* (L.) Moench) Germplasm from Ethiopia and Eritrea: Quantitative Characters. *Euphytica* 115: 91–104. https://doi.org/10.1023/A:1003998313302.

Ayana, Negash Geleta. 2015. Patterns of Variation for Phenotypic Traits in Tetraploid Wheat (*Triticum turgidum* L.) Populations of Ethiopia. *Agricultural and Biological Sciences Journal* 1 (2): 42–51.

Ayer, D.K., A. Sharma, B.R. Ojha, A. Paudel, and K.Dhakal. 2017. Correlation and Path

Coefficient Analysis in Tuberose. *SAAR J.Agri* 15 (1): 1–12.
Azene, Yohannes, Menzir Ahadu, and Tigiest Dejene. 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 and Agriculture* 6: 1–21. https://doi.org/10.1080/23311932.2020.1778604.
Bajgain, P., M. N. Rouse, P. Bulli, S. Bhavani, T. Gordon, R. Wanyera, P. N. Njau, W. Legesse, J. A. Anderson, and M. O. Pumphrey. 2015. Association Mapping of North American Spring Wheat Breeding Germplasm Reveals Loci Conferring Resistance to Ug99 and Other African Stem Rust Races. *BMC Plant Biology* 15 (1): 1–19. https://doi.org/10.1186/s12870-015-0628-9.
Baye, Ashebr, Baye Berihun, Muluken Bantayehu, and Bitwoded Derebe. 2020. Genotypic and Phenotypic Correlation and Path Coefficient Analysis for Yield and Yield-Related Traits in Advanced Bread Wheat (*Triticum aestivum* L.) Lines. *Cogent Food and Agriculture* 6 (1). https://doi.org/10.1080/23311932.2020.1752603.
Bechere, E, G Belay, D Mitiku, and A Merker. 1996. Phenotypic Diversity of Tetraploid Wheat Landraces from Northern and North-Central Regions of Ethiopia. *Hereditas* 124: 165–72.
Bekele, Endashaw. 1984. Analysis of Regional Patterns of Phenotypic Diversity in the Ethiopian Tetraploid and Hexaploid Wheats. *Hereditas* 100 (1): 131–54. https://doi.org/10.1111/j.1601-5223.1984.tb00114.x.
Belay, G, Bechere E., Mitiku D., Merker A., and Tsegaye S. 1997. Patterns of Morphological Diversity in Tetraploid Wheat (*Triticum turgidum* L.) Landraces from Ethiopia. *Acta Agriculturae Scandinavica Section B: Soil and Plant Science* 47: 221–28. https://doi.org/10.1080/09064719709362464.
Belay, Getachew, and Arnulf Merker. 1994. Cytogenetic Studies in Ethiopian Landraces of Tetraploid Wheat (*Triticum turgidum* L.). *Hereditas* 121: 45–52. https://doi.org/10.1111/j.1601-5223.1994.00045.x.
Belay, Getachew, T Tesemma, H C Becker, and A Merker. 1993. Variation and Interrelationships of Agronomic Traits in Ethiopian Tetraploid Wheat Landraces. *Euphytica* 71: 181–88. https://doi.org/10.1007/bf00040407.
Beuselinck, P. R., and J. J. Steiner. 1992. A Proposed Framework for Identifying, Quantifying, and Utilizing Plant Germplasm Resources. *Field Crops Research* 29 (3): 261–72. https://doi.org/10.1016/0378-4290(92)90029-9.
Bhalla, P. L., A. Sharma, and M.B. Singh. 2017. Enabling Molecular Technologies for Trait Improvement in Wheat. In *Wheat Biotechnology*, edited by Prem L Bhalla and Mohan B Singh, 3–24. Springer Sciences Business Media LLC. https://doi.org/10.1007/978-1-4939-7337-8.
Bhalla, Prem, and Mohan B Singh. 2017. *Wheat Biotechnology: Methods and Protocols*.
Bhatta, Madhav, Alexey Morgounov, Vikas Belamkar, and P Stephen Baenziger. 2018. Genome-Wide Association Study Reveals Novel Genomic Regions for Grain Yield and Yield-Related Traits in Drought-Stressed Synthetic Hexaploid Wheat. *International Journal of Molecular Sciences* 19 (3011): 1–23. https://doi.org/10.3390/ijms19103011.
Blackman, J.A. and Payne, P. I.. 1994. Grain Quality. In *WHEAT BREEDING Its Scientific Basis*, edited by G.H. Lupton, F, 455–83.
Blake, Nancy K., Ben R. Lehfeldt, Matt Lavin, and Luther E. Talbert. 1999. Phylogenetic Reconstruction Based on Low Copy DNA Sequence Data in an Allopolyploid: The B Genome of Wheat. *Genome* 42 (2): 351–60. https://doi.org/10.1139/g98-136.
Boucher, Geoff. 2011. Book Reviews: Book Reviews. *Critical Sociology* 37 (4): 493–97. https://doi.org/10.1177/0261018311403863.

Bradbury, Peter J., Zhiwu Zhang, Dallas E. Kroon, Terry M. Casstevens, Yogesh Ramdoss, and Edward S. Buckler. 2007. TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics* 23 (19): 2633–35. https://doi.org/10.1093/bioinformatics/btm308.

Brasesco, F., D. Asgedom, V. Sommacal, and G. Casari. 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. *FAO*. https://doi.org/10.1007/978-94-015-7398-6_11.

Breseghello, Flavio, and Mark E. Sorrells. 2006. Association Mapping of Kernel Size and Milling Quality in Wheat (*Triticum aestivum* L.) Cultivars. *Genetics* 172 (2): 1165–77. https://doi.org/10.1534/genetics.105.044586.

Brett F. Carver. 2009. *Wheat Evolution, Domestication, and Improvement*. Edited by Brett F. Carver. *Wheat : Science and Trade*. Wiley-Blackwell.

Burton, Glenn W, and E H DeVane. 1953. Estimating Heritability in Tall Fescue (*Festzjcu* av-*undinuceu*) from Replicated Clonal Material. *Agronomy Journal*, 478–81.

Bushuk, W. 1998. Wheat Breeding for End-Product Use. *Euphytica* 100 (1–3): 137–45. https://doi.org/10.1023/a:1018368316547.

Bushuk, W., and V.F. Rasper. 1994. *Wheat Production, Properties and Quality*. Edited by W. Bushuk and V. F. Rasper. First edit. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-2672-8.

Caliska, Mahmut. 2012. *GENETIC DIVERSITY*. Edited by Mahmut Çalişkan. inTech. www.intechopen.com.

Cane, Maria Angela, Marco Maccaferri, Ghasemali Nazemi, Silvio Salvi, Rossella Francia, Chiara Colalongo, and Roberto Tuberosa. 2014. Association Mapping for Root Architectural Traits in Durum Wheat Seedlings as Related to Agronomic Performance. *Molecular Breeding* 34 (4): 1629—1645. https://doi.org/10.1007/s11032-014-0177-1.

Chen, Jianhui, Fuyan Zhang, Chunjiang Zhao, Guoguo Lv, Congwei Sun, Yubo Pan, Xinyu Guo, and Feng Chen. 2019. Genome-Wide Association Study of Six Quality Traits Reveals the Association of the TaRPP13L1 Gene with Flour Colour in Chinese Bread Wheat. *Plant Biotechnology Journal* 17: 2106–22. https://doi.org/10.1111/pbi.13126.

Chichti, Emna, Myriam Carrère, Matthieu George, Jean-yves Delenne, Emna Chichti, Myriam Carrère, Matthieu George, Jean-yves Delenne, and Valerie Lullien-pellerin A. 2018. A Wheat Grain Quantitative Evaluation of Vitreousness by Light Transmission Analysis. *Journal of Cereal Science* 83: 58–62. https://doi.org/10.1016/j.jcs.2018.07.009.

Chittaranjan Kole. 2006. *Genome Mapping and Molecular Breeding in Plants: Cereals and Millets*. *Springer*. Vol. 1.

Clercq, Annelies De, and Dirk Inzé. 2006. Cyclin-Dependent Kinase Inhibitors in Yeast, Animals, and Plants: A Functional Comparison. *Critical Reviews in Biochemistry and Molecular Biology* 41 (5): 293–313. https://doi.org/10.1080/10409230600856685.

Cooper, Mark, Fred A Van Eeuwijk, Graeme L Hammer, Dean W Podlich, and Carlos Messina. 2009. Modeling QTL for Complex Traits : Detection and Context for Plant Breeding. *Current Opinion in Plant Biology* 12: 231–40. https://doi.org/10.1016/j.pbi.2009.01.006.

CSA. 2021. The Federa Democratic Republic of Ethiopia Report on Area and Production of Major Crops I.

D. Turner, Stephen. 2018. Qqman: An R Package for Visualizing GWAS Results Using Q-Q and Manhattan Plots. *Journal of Open Source Software* 3 (25): 731.

https://doi.org/10.21105/joss.00731.
Dobraszczyk, B J, M B Whitworth, J F V Vincent, and A A Khan. 2002. Single Kernel Wheat Hardness and Fracture Properties in Relation to Density and the Modelling of Fracture in Wheat Endosperm. *Journal of Cereal Science* 35: 245–63. https://doi.org/10.1006/jcrs.2001.0399.
Doebley, J.F, Gaut. S.B, and D.B. Smith. 2006. The Molecular Genetics of Crop Domestication. *Cell* 45 (5): 1309–21. https://doi.org/10.1016.j.cell.2006.12.006.
Dubcovsky, Jorge, and Jan Dvorak. 2007. "Genome Plasticity a Key Factor." *Science* 316: 1862–66. https://doi.org/10.1126/science.1143986.
Dvorak, J., M. Ch Luo, and E. D. Akhunov. 2011. N.I. Vavilov's Theory of Centres of Diversity in the Light of Current Understanding of Wheat Diversity, Domestication and Evolution. *Czech Journal of Genetics and Plant Breeding* 47 (SPEC. ISSUE 1). https://doi.org/10.17221/3249-cjgpb.
Dvorak, Jan, and Eduard D. Akhunov. 2005. Tempos of Gene Locus Deletions and Duplications and Their Relationship to Recombination Rate during Diploid and Polyploid Evolution in the Aegilops-Triticum Alliance.*Genetics* 171 (1): 323–32. https://doi.org/10.1534/genetics.105.041632.
EAIR. 2017. *Cereals Research Strategy 2016-2030*.
Earl, Dent A., and Bridgett M. vonHoldt. 2012. STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method. *Conservation Genetics Resources* 4 (2): 359–61. https://doi.org/10.1007/s12686-011-9548-7.
Engels, J.M.M. and, J.G. Hawkes, and Melaku Worede. 1991. *Plant Genetic Resources of Ethiopia*. Edited by J.M.K. Engels, J.G. Hawkes, and Melaku Worede. Cambridge University Press. https://doi.org/10.1016/0378-4290(92)90068-k.
Eticha, Firdissa, Endeshaw Bekele, Getachew Belay, and Andreas Börner. 2005. Phenotypic Diversity in Tetraploid Wheats Collected from Bale and Wello Regions of Ethiopia. *Plant Genetic Resources* 3 (1): 35–43. https://doi.org/10.1079/pgr200457.
Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study. *Molecular Ecology* 14 (8): 2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x.
Falush, Daniel, Matthew Stephens, and Jonathan K. Pritchard. 2007. Inference of Population Structure Using Multilocus Genotype Data: Dominant Markers and Null Alleles. *Molecular Ecology Notes* 7 (4): 574–78. https://doi.org/10.1111/j.1471-8286.2007.01758.x.
FAO. 2011. *Save and Grow: A Policymaker's Guide to the Sustainable Intensfication of Smallholder Crop Production*. Rome.
———. 2017. The Future of Food and Agriculture: Trends and Challenges. *FAO*. Vol. 4. Rome.
———. 2019. *World Food and Agriculture - Statistical Pocketbook 2019*. *World Food and Agriculture - Statistical Pocketbook 2020*. Rome. https://doi.org/10.4060/cb1521en.
Feldman, Moshe, and Mordechai E. Kislev. 2007. Domestication of Emmer Wheat and Evolution of Free-Threshing Tetraploid Wheat.*Israel Journal of Plant Sciences* 55 (3–4): 207–21. https://doi.org/10.1560/IJPS.55.3-4.207.
Feldman, Moshe, and A.Avraham Levy. 2015. Origin and Evolution of Wheat and Related Triticeae Species. In *Alien Introgression in Wheat: Cytogenetics, Molecular Biology, and Genomics*, edited by Marta Molnar_Lang, Carla and Ceoloni, and Jaroslav Dolezel, 21–76. Springer International Publishing. https://doi.org/10.1007/978-3-319-23494-6.
Flint-garcia, Sherry A, Jeffry M Thornsberry, and Edward S Buckler Iv. 2003. Structure of

Linkage Disequilibrium in Plants. *Annual Review Plan Biology*. https://doi.org/10.1146/annurev.arplant.54.031902.134907.

Frankham, R., J.D. and Ballou, and D.A. Briscoe. 2002. *Introduction to Conservation Genetics*. Cambridge University Press.

Fufa, Miheretu, Andargachew Gedebo, Tesfaye Letta, and Dangachew Lule. 2023. Genetic Diversity and Population Structure Analysis in Tetraploid Wheat (*Triticum turgidum* Spp) Germplasm from Ethiopia Based on DArTSeq Markers. *Genetic Resources and Crop Evolution* 71 (6): 2415–33. https://doi.org/10.1007/s10722-023-01775-6.

Ganno, James, Dawit Alemu, and Gidada Ayalew. 2017. A Study of Genetic Variation and Grain Quality Traits in Bread Wheat (*Tritium aestivum* L.) Genotypes. *African Journal of Plant Breeding* 4 (1): 172–82.

Gao, Le, Chengsheng Meng, Tengfei Yi, Ke Xu, Huiwen Cao, Shuhua Zhang, and Xueju Yang. 2021. Genome-Wide Association Study Reveals the Genetic Basis of Yield- and Quality-Related Traits in Wheat. *BMC Plant Biology* 21 (144): 1–11. https://doi.org/https://doi.org/10.1186/s12870-021-02925-7.

Gashaw, A, H Mohammed, and H Singh. 2010. Genetic Divergence in Selected Durum Wheat Genotypes of Ethiopian Plasm. *African Crop Science Journal* 15 (2): 67–72. https://doi.org/10.4314/acsj.v15i2.54419.

Gebreselassie, Samuel, Mekbib G. Haile, and Matthias Kalkuhl. 2017. The Wheat Sector in Ethiopia: Current Status and Key Challenges for Future Value Chain Development. Addis Abeba.

Geleta, N, and H Grausgruber. 2014. Homemade Products and Socio-Cultural Values of Wheat Seed Production in Ambo and Dandi Districts of West Central Ethiopia. *Science, Technology and Arts Research Journal* 2 (4): 62. https://doi.org/10.4314/star.v2i4.11.

Geleta, Negash, and Heinrich Grausgruber. 2012. Classifying Ethiopian Tetraploid Wheat (*Triticum turgidum* L.) Landraces by Combined Analysis of Molecular & Phenotypic Data. *Science, Technology and Arts Research Journal* 1 (1): 1–9. https://doi.org/10.4314/star.v1i1.98767.

———. 2016. Morphological and Quality Traits Variation in Tetraploid (Triticum Turgidum l.) and Hexaploid (*Triticum aestivum* l.) Wheat Accessions from Ethiopia. *Agricultural Science Research Journals* 3 (8): 229–36.

Getachew, Ahmed, Sentayehu Alamerew, and Fisseha Worede. 2017. Multivariate Analyses of Phenotypic Diversity of Bread Wheat (*Triticum aestivum* L.) in the Highlands of Northern Ethiopia. *Advances in Crop Science and Technology* 5 (5): 1–7. https://doi.org/10.4172/2329-8863.1000309.

Getachew, T Tesfaye B, and M Worede. 1991. Morphological Diversity in Tetraploid Wheat Landrace Popula- Tions from the Central Highlands of Ethiopia. *Hereditas* 114: 171–76.

Gibbon, David. 2012. Save and Grow: A Policymaker's Guide to the Sustainable Intensification of Smallholder Crop Production. Rome, Italy: Food and Agriculture Organization of the United Nations (2011), Pp. 112. *Experimental Agriculture* 48 (1): 154–154. https://doi.org/10.1017/s0014479711001049.

Giraldo, Patricia, Conxita Royo, Mirvana González, Jose M Carrillo, and Magdalena Ruiz. 2016. *Genetic Diversity and Association Mapping for Agromorphological and Grain Quality Traits of a Structured Collection of Durum Wheat Landraces Including Subsp. Durum, Turgidum and Diccocon*. *PLoS ONE*. Vol. 11. https://doi.org/10.1371/journal.pone.0166577.

Gomez, Kwanchai A, and Arturo A Gomez. 1984. *Statistical Procedures for Agricultural Research*.

González-Diéguez, Jose Crossa, Carolina Rivera-amado, Francisco Pinera-chavez, Carolina Saint Pierre, and David González Diéguez. 2022. GENOME-BASED PREDICTIONS OF SUB-GENOME GENETIC INTERACTIONS EFFECTS IN WHEAT Evolutionary History of Common Wheat.*XVIIIth Eucarpia Biometrics in Plant Breeding Conference. CYMMIT*.

Gopher, Avi, Shahai and Abbo, and Simcha Lev-Yadun. 2000. Documenta Praehistorica 15: 49–62.

Govindaraj, M., M. Vetriventhan, and M. Srinivasan. 2015. Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. *Genetics Research International* 2015: 1–14. https://doi.org/10.1155/2015/431487.

Grenier, C., P. J. Bramel, J. A. Dahlberg, A. El-Ahmadi, M. Mahmoud, G. C. Peterson, D. T. Rosenow, and G. Ejeta. 2004. Sorghums of the Sudan: Analysis of Regional Diversity and Distribution. *Genetic Resources and Crop Evolution* 51 (5): 489–500. https://doi.org/10.1023/B:GRES.0000024155.43149.71.

Gunupuru, Lokanadha R., Chanemougasoundharam Arunachalam, Keshav B. Malla, Amal Kahla, Alexandre Perochon, Jianguang Jia, Ganesh Thapa, and Fiona M. Doohan. 2018. A Wheat Cytochrome P450 Enhances Both Resistance to Deoxynivalenol and Grain Yield. *PLoS ONE* 13 (10): 1–17. https://doi.org/10.1371/journal.pone.0204992.

Gupta, P K, Jitendra Kumar, R R Mir, and Ajay Kumar. 2010. Marker-Assisted Selection as a Component of Conventional. *Plant Breeding Reviews* 33: 145–217.

Hagenblad, Jenny, Jenny Zie, and Matti W. Leino. 2012. Exploring the Population Genetics of Genebank and Historical Landrace Varieties. *Genetic Resources and Crop Evolution* 59 (6): 1185–99. https://doi.org/10.1007/s10722-011-9754-x.

Haile, Jemanesh, Karl Hammer, Badebo Ayele, Miloudi M. Nachit, and Marion S. Röder. 2013. Genetic Diversity Assessment of Ethiopian Tetraploid Wheat Landraces and Improved Durum Wheat Varieties Using Microsatellites and Markers Linked with Stem Rust Resistance. *Genetic Resources and Crop Evolution* 60: 513–27. https://doi.org/10.1007/s10722-012-9855-1.

Haile, Jemanesh, Karl Hammer, Ayele Badebo, Ravi P. Singh, and Marion S. Roder. 2013. Haplotype Analysis of Molecular Markers Linked to Stem Rust Resistance Genes in Ethiopian Improved Durum Wheat Varieties and Tetraploid Wheat Landraces. *Genetic Resources and Crop Evolution* 60: 853–64. https://doi.org/10.1007/s10722-012-9880-0.

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

Hailu, Faris, Marker A, Belay G, and Eva Johansson. 2005. Molecular Diversity and Phylogenic Relationships of Tetraploid Wheat Species as Revealed by Inter-Simple Sequence Repeats (ISSR) from Ethiopia. *Journal of Genetics and Breeding* 59: 1–9.

Hailu, Faris, Eva Johansson, and Arnulf Merker. 2010. Patterns of Phenotypic Diversity for Phenologic and Qualitative Traits in Ethiopian Tetraploid Wheat Germplasm. *Genetic Resources and Crop Evolution* 57: 781–90. https://doi.org/10.1007/s10722-009-9518-z.

Hailu, Faris, Eva Johansson, Arnulf Merker, Getachew Belay, Harjit Singh, and Habtamu Zeleke. 2006. *Composition of and Variation in High- and Low-Molecular Weight Glutenin Subunits, and Omega Gliadins in Ethiopian Tetraploid Wheat Germplasm*. *Plant Genetic Resources*. Vol. 4. HINARI. https://doi.org/10.1079/pgr2006110.

Hailu, Faris, and Arnulf Merker. 2008. Variation in Gluten Strength and Yellow Pigment in Ethiopian Tetraploid Wheat Germplasm. *Genetic Resources and Crop Evolution*

55: 277–85. https://doi.org/10.1007/s10722-007-9233-6.
Hailu, Faris, Arnulf Merker, Belay and Getachew, and Eva Johanssonx. 2006. Multivariate Analysis of Diversity of Tetraploid Wheat Germplasm from Ethiopia. *Genetic Resources and Crop Evolution* 53: 1089–98. https://doi.org/10.1007/s10722-005-9776-3.
Halder, Jyotirmoy, Harsimardeep S. Gill, Jinfeng Zhang, Rami Altameemi, Eric Olson, Brent Turnipseed, and Sunish K. Sehgal. 2023. Genome-Wide Association Analysis of Spike and Kernel Traits in the U.S. Hard Winter Wheat. *Plant Genome* 16 (1): 1–18. https://doi.org/10.1002/tpg2.20300.
Hardon, J.J. 1992. Plant Genetic Resources of Ethiopia. *Field Crops Research* 30 (1–2): 187–88. https://doi.org/10.1016/0378-4290(92)90068-k.
Harjes, Carlos E. *et al.* 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification, *Science*, 319(5861), pp. 330–333. https://doi.org/10.1126/science.1150255.

Harlan, J.R., J.M.J. De Wet, and E.G Price. 1973. Comparative Evolution of Cereals. *Evolution* 27 (2): 311–25. https://doi.org/doi: 10.1111/j.1558-5646.1973.tb00676.x.

Harlan, Jack R. 1969. Ethiopia: A Center of Diversity. *Economic Botany* 23 (4): 309–14. https://doi.org/10.1007/BF02860676.
Harlan, Jack R., and Daniel Zohary. 1966. Distribution of Wild Wheats and Barley. *Science* 153 (3740): 1074–80. https://doi.org/10.1126/science.153.3740.1074.
Harris, Richard. 2018. An Introduction to R. *Quantitative Geography: The Basics* 1: 250–86. https://doi.org/10.4135/9781473920446.n12.
Hartl, Daniel L., and Andreq G. Clark. 1997. *Principles of Population Genetics*. 3rd ed. Sinauer Associates, Inc. Publishers. https://doi.org/10.1111/j.1439-0388.2000.201-2.x.
———. 2007. *Principles of Population Genetics, 4th Edition*. *Journal of Heredity*. Fourth. Vol. 98. Sinauer Associates, Inc. Publishers. https://doi.org/10.1093/jhered/esm035.
Hartley, H. O. 1950. The Maximum F-Ratio as a Short-Cut Test for Heterogeneity of Variance. *Biometrika* 37 (3/4): 308. https://doi.org/10.2307/2332383.
Heun, Manfred, Ralf Schäfer-Pregl, Dieter Klawan, Renato Castagna, Monica Accerbi, Basilio Borghi, and Francesco Salamini. 1997. Site of Einkorn Wheat Domestication Identified by DNA Fingerprinting. *Science* 278 (5341): 1312–14. https://doi.org/10.1126/science.278.5341.1312.
Hill, W. G., and B. S. Weir. 1988. Variances and Covariances of Squared Linkage Disequilibria in Finite Populations. *Theoretical Population Biology* 33 (1): 54–78. https://doi.org/10.1016/0040-5809(88)90004-4.
Hirschhorn, Joel N., and Mark J. Daly. 2005. Genome-Wide Association Studies for Common Diseases and Complex Traits. *Nature Reviews Genetics* 6 (2): 95–108. https://doi.org/10.1038/nrg1521.
https://www.kddart.org/kdcompute.html
https://kdcompute.seqart.net/kdcompute/plugins
https://urgi.versailles.inra.fr/download/iwgsc/
https://wheat-urgi.versailles,inrae.fr/Seq-Repository/Annotations
https://statisticalpoint.com/anova-f-value-p-value
https://www.mdpi.com/ijms/ijms-21-05836/article_deploy/html/images/ijms-21-05836-g001.png
Hu, Ming Jian, Hai Ping Zhang, Kai Liu, Jia Jia Cao, Sheng Xing Wang, Hao Jiang, Zeng Yun Wu, et al. 2016. Cloning and Characterization of TaTGW-7A Gene Associated with Grain Weight in Wheat via SLAF-Seq-BSA. *Frontiers in Plant Science* 7 (DECEMBER2016): 1–12. https://doi.org/10.3389/fpls.2016.01902.

Hu, Xin, Jing Ren, Xifeng Ren, Sisi Huang, Salih A.I. Sabiel, Mingcheng Luo, Eviatar Nevo, Chunjie Fu, Junhua Peng, and Dongfa Sun. 2015. Association of Agronomic Traits with SNP Markers in Durum Wheat (*Triticum turgidum* L. Durum (Desf.)). *PLoS ONE* 10 (6): 1–23. https://doi.org/10.1371/journal.pone.0130854.

Huang, Meng, Xiaolei Liu, Yao Zhou, Ryan M. Summers, and Zhiwu Zhang. 2019. BLINK: A Package for the next Level of Genome-Wide Association Studies with Both Individuals and Markers in the Millions. *GigaScience* 8 (2): 1–12. https://doi.org/10.1093/gigascience/giy154.

Huang, Shaoxing, Anchalee Sirikhachornkit, Xiujuan Su, Justin Faris, Bikram Gill, Robert Haselkorn, and Piotr Gornicki. 2002. Genes Encoding Plastid Acetyl-CoA Carboxylase and 3-Phosphoglycerate Kinase of the Triticum/Aegilops Complex and the Evolutionary History of Polyploid Wheat. *Proceedings of the National Academy of Sciences of the United States of America* 99 (12): 8133–38. https://doi.org/10.1073/pnas.072223799.

Huang, X. Q., A. Börner, M. S. Röder, and M. W. Ganal. 2002. Assessing Genetic Diversity of Wheat (*Triticum aestivum* L.) Germplasm Using Microsatellite Markers. *Theoretical and Applied Genetics* 105 (5): 699–707. https://doi.org/10.1007/s00122-002-0959-4.

IBC. 2012. A Guide to Access to Genetic Resources and Community Knowledge and Benefit Sharing in Ethiopia.

IBPGR. 1985. Descriptors_Wheat.Pdf.

Inda, Luis A., José Gabriel Segarra-Moragues, Jochen Müller, Paul M. Peterson, and Pilar Catalán. 2008. Dated Historical Biogeography of the Temperate Loliinae (Poaceae, Pooideae) Grasses in the Northern and Southern Hemispheres. *Molecular Phylogenetics and Evolution* 46 (3): 932–57. https://doi.org/10.1016/j.ympev.2007.11.022.

Jaccoud, Damian, Kaiman Peng, David Feinstein, and Andrzej Kilian. 2001. Diversity Arrays: A Solid State Technology for Sequence Information Independent Genotyping. *Nucleic Acids Research* 29 (4): 25. https://doi.org/10.1093/nar/29.4.e25.

Jain, Mukesh, Aashima Nijhawan, Rita Arora, Pinky Agarwal, Swatismita Ray, Pooja Sharma, Sanjay Kapoor, Akhilesh K. Tyagi, and Jitendra P. Khurana. 2007. F-Box Proteins in Rice. Genome-Wide Analysis, Classification, Temporal and Spatial Gene Expression during Panicle and Seed Development, and Regulation by Light and Abiotic Stress. *Plant Physiology* 143 (4): 1467–83. https://doi.org/10.1104/pp.106.091900.

Jain, S. K., C. O. Qualset, G. M. Bhatt, and K. K. Wu. 1975. Geographical Patterns of Phenotypic Diversity in a World Collection of Durum Wheats 1 . *Crop Science* 15 (5): 700–704. https://doi.org/10.2135/cropsci1975.0011183x001500050026x.

Jamil, Muhammad, Aamir Ali, Alvina Gul, Abdul Ghafoor, Abdul Aziz Napar, Amir M H Ibrahim, Naima Huma Naveed, Nasim Ahmad Yasin, and Abdul Mujeeb-kazi. 2019. Genome-Wide Association Studies of Seven Agronomic Traits under Two Sowing Conditions in Bread Wheat. *BMC Plant Biology* 19 (149): 1–18. https://doi.org/https://doi.org/10.1186/s12870-019-1754-6.

Jantasuriyarat, C., M. I. Vales, C. J.W. Watson, and O. Riera-Lizarazu. 2004. Identification and Mapping of Genetic Loci Affecting the Free-Threshing Habit and Spike Compactness in Wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 108 (2): 261–73. https://doi.org/10.1007/s00122-003-1432-8.

Johnson, Herbert W, H.F. Robinson, and R.E.Comstock. 1955. Estimates of Genetic and Environmental Variability in Sesame. *Agronomy*, 314–18. https://doi.org/10.1017/S0014479700009510.

Johnson, Marina, and Marina Johnson. 2017. Association Mapping and Genetic Diversity Studies of Agronomic and Quality Traits in Durum Wheat [*Triticum turgidum* L. Var. Durum (Desf.)]. https://library.ndsu.edu/ir/handle/10365/27462.

Johnson, Marina, Ajay Kumar, Atena Oladzad-Abbasabadi, Evan Salsman, Meriem Aoun, Frank A. Manthey, and Elias M. Elias. 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: 1–18. https://doi.org/10.3389/fgene.2019.00717.

Johson, Richard A., and Dean W. Wichiern. 2007. *Applied Multivariate Statistical Analysis*. *Pearson Education*.

Joshi, Anjali, Anil Kumar, Vartika Budhlakoti, Nidhi Bhatt And, and Tabassum. 2018. Estimation of Genetic Variability Parameters for Yield and Its Components in Bread Wheat (*Triticum aestivum*). *Advances in Agriculture and Natural Sciences for Sustainable Agriculture*, no. 5: 87–90.

Jung, Woo Joo, Yong Jin Lee, Chon-sik Kang, and Yong Weon Seo. 2021. Identification of Genetic Loci Associated with Major Agronomic Traits of Wheat (*Triticum aestivum* L .) Based on Genome- Wide Association Analysis." *BMC Plant Biology* 21 (418): 1–14. https://doi.org/https://doi.org/10.1186/s12870-021-03180-6.

Kabbaj, Hafssa, Amadou T Sall, Ayed Al-abdallat, Mulatu Geleta, Ahmed Amri, and Bari Aldo. 2017. Genetic Diversity within a Global Panel of Durum Wheat (*Triticum durum*) Landraces and Modern Germplasm Reveals the History of Alleles Exchange." *Frontiers in Plant Science* 8 (July): 1–13. https://doi.org/10.3389/fpls.2017.01277.

Kadkol, G P, and M Sissons. 2016. *Durum Wheat : Overview*. *Encyclopedia of Food Grains*. 2nd ed. Elsevier Ltd. https://doi.org/10.1016/B978-0-12-394437-5.00024-3.

Kearsey, M.J, and H.S. Pooni. 1996. *The Genetical Analysis of Quantitative Traits*. *Journal of Medical Genetics*. Vol. 33. https://doi.org/10.1136/jmg.33.11.976.

Keegstra, Kenneth, and Natasha Raikhel. 2001. Plant Glycosyltransferases. *Current Opinion in Plant Biology* 4 (3): 219–24. https://doi.org/DOI: 10.1016/s1369-5266(00)00164-3.

Keler, V. V., O. V. Martynova, and T. G. Ovchinnikova. 2020. Pesticides Effect on the Quantity and Quality of Gluten in Spring Wheat. In *IOP Conference Series: Earth and Environmental Science*, 548:1–7. IOP Publishing Ltd. https://doi.org/10.1088/1755-1315/548/5/052036.

Kellogg, Elizabeth a. 2001. Update on Evolution Evolutionary History of the Grasses 1. *Plant Physiology* 125: 1198–1205. http://www.plantphysiol.org.

Khorgade, P., W. Narkhede and S.K. Raut, 1985. 1985. Genetic Variability in Chickpea. *International Chickpea Newsletter*, 1985.

Kidane, Yosef G., Bogale N. Hailemariam, Dejene K. Mengistu, Carlo Fadda, Mario E. and Pe, and Matteo Dell'Acqua. 2017. Genome-Wide Association Study of Septoria Tritici Blotch Resistance in Ethiopian Durum Wheat Landraces. *Frontiers in Genetics* 8: 1–12. https://doi.org/10.3389/fpls.2017.01586.

Kidane, Yosef G, Cherinet A Gesesse, Bogale N Hailemariam, Ermias A Desta, Dejene K Mengistu, Carlo Fadda, and Mario Enrico P. 2019. A Large Nested Association Mapping Population for Breeding and Quantitative Trait Locus Mapping in Ethiopian Durum Wheat. *Plant Biotechnology Journal* 17: 1380–93. https://doi.org/10.1111/pbi.13062.

Kidane, Yosef G, Chiara Mancini, Dejene K Mengistu, Elisabetta Frascaroli, Carlo Fadda, Mario Enrico Pè, and Matteo Dell'Acqua. 2017. Genome Wide Association Study to Identify the Genetic Base of Smallholder Farmer Preferences of Durum Wheat Traits.

Frontiers in Plant Sciencerontiers in Plant Science 8: 1–11. https://doi.org/10.3389/fpls.2017.01230.

Kilian, Andrzej, Peter Wenzl, Eric Huttner, Jason Carling, Ling Xia, Hélène Blois, Vanessa Caig, et al. 2012. Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms. *Methods in Molecular Biology* 888: 67–89. https://doi.org/10.1007/978-1-61779-870-2_5.

Kilian, Benjamin, William and Martin, and Francesco Salamini. 2010. Evolution in Action: Case Studies in Adaptive Radiation, Speciation and the Origin of Biodiversity. *Evolution in Action: Case Studies in Adaptive Radiation, Speciation and the Origin of Biodiversity*, 1–586. https://doi.org/10.1007/978-3-642-12425-9.

Kong, Lingan, Jisheng Si, Bin Zhang, Bo Feng, Shengdong Li, and Fahong Wang. 2013. Environmental Modification of Wheat Grain Protein Accumulation and Associated Processing Quality: A Case Study of China. *Australian Journal of Crop Science* 7 (2): 173–81.

Korkut, Zahit, O B N Lg, and Nezihi S A Lam. 2007. Stability of Grain Vitreousness in Durum Wheat (*Triticum durum* L. Desf.) Genotypes in the North-Western Region of Turkey. *Turkish Journal of Agriculture and Forestry* 31: 313–18.

Korte, Arthur, and Ashley Farlow. 2013. The Advantages and Limitations of Trait Analysis with GWAS : A Review. *Plant Methods*, 1–9. https://doi.org/10.1186/1746-4811-9-29.

Kraakman, Arnold T.W., Rients E. Niks, Petra M.M.M. Van Den Berg, Piet Stam, and Fred A. Van Eeuwijk. 2004. Linkage Disequilibrium Mapping of Yield and Yield Stability in Modern Spring Barley Cultivars. *Genetics* 168 (1): 435–46. https://doi.org/10.1534/genetics.104.026831.

Kumar, Deepender, Vinod Chhokar, Sonia Sheoran, Rajender Singh, Pradeep Sharma, and Sarika Jaiswal. 2019. Characterization of Genetic Diversity and Population Structure in Wheat Using Array Based SNP Markers. *Molecular Biology Reports*, no. 0123456789. https://doi.org/10.1007/s11033-019-05132-8.

Kumar, Narender, and John C. Larkin. 2017. Why Do Plants Need so Many Cyclin-Dependent Kinase Inhibitors? *Plant Signaling and Behavior* 12 (2): 1–4. https://doi.org/10.1080/15592324.2017.1282021.

Kumar, Ravindra. 2020. Genetic Variability among Wheat Genotypes Based on Agro-Morphological Traits. *Phytojournal.Com* 9: 57–64. http://www.phytojournal.com/archives/2020/vol9issue3/PartM/9-2-398-659.pdf.

Kumar, Sundeep, Jyoti Kumari, Nabin Bhusal, Anjan Kumar Pradhan, Neeraj Budhlakoti, Dwijesh Chandra Mishra, Divya Chauhan, et al. 2020. Genome-Wide Association Study Reveals Genomic Regions Associated With Ten Agronomical Traits in Wheat Under Late-Sown Conditions. *Frontiers in Plant Science* 11 (Article 549743): 1–15. https://doi.org/10.3389/fpls.2020.549743.

Laidò, Giovanni, Giacomo Mangini, Francesca Taranto, Agata Gadaleta, Antonio Blanco, Luigi Cattivelli, Daniela Marone, Anna M. Mastrangelo, Roberto Papa, and Pasquale De Vita. 2013. Genetic Diversity and Population Structure of Tetraploid Wheats (*Triticum turgidum* L.) Estimated by SSR, DArT and Pedigree Data. *PLoS ONE* 8 (6): 1–17. https://doi.org/10.1371/journal.pone.0067280.

Letta, Tesfaye, Pablo Olivera, Marco Maccaferri, Yue Jin, Karim Ammar, Ayele Badebo, Silvio Salvi, Enrico Noli, Jose Crossa, and Roberto Tuberosa. 2014. Association Mapping Reveals Novel Stem Rust Resistance Loci in Durum Wheat at the Seedling Stage. *The Plant Genome* 7 (1). https://doi.org/10.3835/plantgenome2013.08.0026.

Li, Penghui, Enhua Xia, Jiamin Fu, Yujie Xu, Xuecheng Zhao, Wei Tong, Qian Tang, Million Tadege, Alisdair R. Fernie, and Jian Zhao. 2022. Diverse Roles of MYB

Transcription Factors in Regulating Secondary Metabolite Biosynthesis, Shoot Development, and Stress Responses in Tea Plants (*Camellia sinensis*). *Plant Journal* 110 (4): 1144–65. https://doi.org/10.1111/tpj.15729.

Li, S, L Xi, H Liu, W Zhu, L Xu, Y Wang, J Zeng, X Fan, and L Sha. 2020. The Classification of Tetraploid Wheat by Phylogenetic and Cytogenetic Analyses. *Biologia Plantarum* 64: 324–34. https://doi.org/10.32615/bp.2020.013.

Li, Weixing, Siyu Pang, Zhaogeng Lu, and Biao Jin. 2020. Function and Mechanism of WRKY Transcription Factors in Abiotic Stress Responses of Plants. *Plants* 9 (11): 1–15. https://doi.org/10.3390/plants9111515.

Li, Yixuan, and Kaifa Wei. 2020. Comparative Functional Genomics Analysis of Cytochrome P450 Gene Superfamily in Wheat and Maize. *BMC Plant Biology* 20 (1): 1–23. https://doi.org/10.1186/s12870-020-2288-7.

Lim, Eng-kiat, and Dianna J Bowles. 2004. A Class of Plant Glycosyltransferases Involved in Cellular Homeostasis. *The EMBO Journal* 23 (15): 2915–22. https://doi.org/10.1038/sj.emboj.7600295.

Lipka, Alexander E., Feng Tian, Qishan Wang, Jason Peiffer, Meng Li, Peter J. Bradbury, Michael A. Gore, Edward S. Buckler, and Zhiwu Zhang. 2012. GAPIT: Genome Association and Prediction Integrated Tool. *Bioinformatics* 28 (18): 2397–99. https://doi.org/10.1093/bioinformatics/bts444.

Liu, Guozheng, Yusheng Zhao, Vilson Mirdita, and Jochen Christoph Reif. 2017. Efficient Strategies to Assess Yield Stability in Winter Wheat. *Theoretical and Applied Genetics* 130 (8): 1587–99. https://doi.org/10.1007/s00122-017-2912-6.

Liu, Hang, Jian Ma, Yang Tu, Jing Zhu, Puyang Ding, Jiajun Liu, Ting Li, et al. 2020. Several Stably Expressed QTL for Spike Density of Common Wheat (*Triticum aestivum*) in Multiple Environments. *Plant Breeding* 139 (2): 284–94. https://doi.org/10.1111/pbr.12782.

Liu, Kejun, and Spencer V. Muse. 2005. PowerMaker: An Integrated Analysis Environment for Genetic Maker Analysis. *Bioinformatics* 21 (9): 2128–29. https://doi.org/10.1093/bioinformatics/bti282.

Liu, Weizhen, Marco Maccaferri, Xianming Chen, and Gaetano Laghetti. 2017. Genome - Wide Association Mapping Reveals a Rich Genetic Architecture of Stripe Rust Resistance Loci in Emmer Wheat (*Triticum turgidum* Ssp . Dicoccum). *Theoretical and Applied Genetics* 130 (11): 2249–70. https://doi.org/10.1007/s00122-017-2957-6.

Liu, Weizhen, Marco Maccaferri, Sheri Rynearson, Tesfaye Letta, Habtemariam Zegeye, Roberto Tuberosa, Xianming Chen, and Michael Pumphrey. 2017. Novel Sources of Stripe Rust Resistance Identified by Genome-Wide Association Mapping in Ethiopian Durum Wheat (*Triticum turgidum* Ssp. Durum). *Frontiers in Plant Science* 8 (Article 774): 1–15. https://doi.org/10.3389/fpls.2017.00774.

Lopes, M. S., S. Dreisigacker, R. J. Peña, S. Sukumaran, and M. P. Reynolds. 2014. Genetic Characterization of the Wheat Association Mapping Initiative (WAMI) Panel for Dissection of Complex Traits in Spring Wheat. *Theoretical and Applied Genetics*. https://doi.org/10.1007/s00122-014-2444-2.

Lozada, Dennis N., R. Esten Mason, Md Ali Babar, Brett F. Carver, Gina Brown Guedira, Keith Merrill, Maria Nelly Arguello, et al. 2017. Association Mapping Reveals Loci Associated with Multiple Traits That Affect Grain Yield and Adaptation in Soft Winter Wheat. *Euphytica* 213 (9): 1–15. https://doi.org/10.1007/s10681-017-2005-2.

Maccaferri, M., M. C. Sanguineti, P. Donini, and R. Tuberosa. 2003. Microsatellite Analysis Reveals a Progressive Widening of the Genetic Basis in the Elite Durum Wheat Germplasm" *Theoretical and Applied Genetics* 107 (5): 783–97. https://doi.org/10.1007/s00122-003-1319-8.

Maccaferri, Marco, Maria C Sanguineti, Paola Mantovani, Jerzy H Czembor, Smadar and Ezrati, and Roberto Tuberosa. 2010. Association Mapping of Leaf Rust Response in Durum Wheat. *Molecular Breeding* 26: 189–228. https://doi.org/10.1007/s11032-009-9353-0.

Maccaferri, Marco, Maria Corinna Sanguineti, Simona Corneti, Jose Luis Araus Ortega, Moncef Ben Salem, Jordi Bort, Enzo Deambrogio, et al. 2008. Quantitative Trait Loci for Grain Yield and Adaptation of Durum Wheat (*Triticum durum* Desf.) Across a Wide Range of Water Availability. *Genetics* 178: 489–511. https://doi.org/10.1534/genetics.107.077297.

MacCaferri, Marco, Maria Corinna Sanguineti, Andrea Demontis, Ahmed El-Ahmed, Luis Garcia Del Moral, Fouad Maalouf, Miloudi Nachit, et al. 2011. Association Mapping in Durum Wheat Grown across a Broad Range of Water Regimes. *Journal of Experimental Botany* 62 (2): 409–38. https://doi.org/10.1093/jxb/erq287.

Mackay, Trudy F C, Eric A Stone, and Julien F Ayroles. 2009. The Genetics of Quantitative Traits : Challenges and Prospects. *Nature Reviews Genetics* 10. https://doi.org/10.1038/nrg2612.

Mammadov, Jafar, Rajat Aggarwal, Ramesh Buyyarapu, and Siva Kumpatla. 2012. SNP Markers and Their Impact on Plant Breeding. https://doi.org/10.1155/2012/728398.

Marone, Daniela, Giosuè Panio, Donatella B.M. Ficco, Maria A. Russo, Pasquale De Vita, Roberto Papa, Diego Rubiales, Luigi Cattivelli, and Anna M. Mastrangelo. 2012. Characterization of Wheat DArT Markers: Genetic and Functional Features. *Molecular Genetics and Genomics*, 1–13. https://doi.org/10.1007/s00438-012-0714-8.

Mathew, Isack, Shimelis Hussein, Admire Isaac, Tichafa Shayanowako, Mark Laing, and Vincent Chaplot. 2019. Genome-Wide Association Study of Drought Tolerance and Biomass Allocation in Wheat. *PLoS ONE* 14 (12): 1–21. https://doi.org/https://doi.org/10.1371/journal. pone.0225383.

Mauricio, Rodney. 2001. MAPPING QUANTITATIVE TRAIT LOCI IN PLANTS : USES AND CAVEATS FOR EVOLUTIONARY BIOLOGY. *Nature Reviews Genetics* 2: 370–81.

Mehrabi, Ali Ashraf, Alireza Pour-Aboughadareh, Sajjad Mansouri, and Alireza Hosseini. 2020. Genome-Wide Association Analysis of Root System Architecture Features and Agronomic Traits in Durum Wheat. *Molecular Breeding* 40 (6). https://doi.org/10.1007/s11032-020-01136-6.

Meirmans, Patrick G., and Shenglin Liu. 2018. Analysis of Molecular Variance (AMOVA) for Autopolyploids. *Frontiers in Ecology and Evolution* 6 (MAY): 1–10. https://doi.org/10.3389/fevo.2018.00066.

Mekonnen, Molla, Getenet Sharie, Muluken Bayable, Amleku Teshager, Eshetie Abebe, Misganaw Ferede, Desalew Fentie, et al. 2020. Participatory Variety Selection and Stability Analysis of Durum Wheat Varieties (*Triticum durum* Desf.) in Northwest Amhara. *Cogent Food and Agriculture* 6 (1). https://doi.org/10.1080/23311932.2020.1746229.

Meles, Berhanu, Wassu Mohammed, and Yemane Tsehaye. 2017. Genetic Variability , Correlation and Path Analysis of Yield and Grain Quality Traits in Bread Wheat (*Tritium aestivum* L.) Genotypes at Axum , Northern Ethiopia" 9 (October): 175–85. https://doi.org/10.5897/JPBCS2016.0671.

Mengistu, Dejene K., Afeworki Y. Kiros, and Mario E. Pè. 2015a. *Phenotypic Diversity in Ethiopian Durum Wheat (Triticum Turgidum Var. Durum) Landraces*. *Crop Journal*. Vol. 3. https://doi.org/10.1016/j.cj.2015.04.003.

Mengistu, Dejene K, Yosef G Kidane, Carlo Fadda, and Mario Enrico Pè. 2016. Genetic Diversity in Ethiopian Durum Wheat (*Triticum turgidum* Var Durum) Inferred from

Phenotypic Variations. *Plant Genetic Resources*, 1–11. https://doi.org/10.1017/S1479262116000393.

Mengistu, Dejene K, Afeworki Y Kiros, and Mario E Pè. 2015b. Phenotypic Diversity in Ethiopian Durum Wheat (*Triticum turgidum* Var. Durum) Landraces. *Crop Journal* 3: 190–99. https://doi.org/10.1016/j.cj.2015.04.003.

Mengistu, Dejene K, and Mario E Pè. 2016. Revisiting the Ignored Ethiopian Durum Wheat (*Triticum turgidum* Var . Durum) Landraces for Genetic Diversity Exploitation in Future Wheat Breeding Programs" 8 (4): 45–59. https://doi.org/10.5897/JPBCS2015.0542.

Mengistu, Dejene Kassahun, Yosef Gebrehawaryat Kidane, Marcello Catellani, Elisabetta Frascaroli, Carlo Fadda, Mario Enrico Pè, and Matteo Dell'Acqua. 2016. High-Density Molecular Characterization and Association Mapping in Ethiopian Durum Wheat Landraces Reveals High Diversity and Potential for Wheat Breeding. *Plant Biotechnology Journal* 14: 1800–1812. https://doi.org/10.1111/pbi.12538.

Mérida-García, Rosa, Guozheng Liu, Sang He, Victoria Gonzalez-Dugo, Gabriel Dorado, Sergio Gálvez, Ignacio Solís, Pablo J. Zarco-Tejada, Jochen C. Reif, and Pilar Hernandez. 2019. Genetic Dissection of Agronomic and Quality Traits Based on Association Mapping and Genomic Selection Approaches in Durum Wheat Grown in Southern Spain. *PLoS ONE* 14 (2): 1–24. https://doi.org/10.1371/journal.pone.0211718.

Mesele, Adhiena, Mohammed Wassu, and Dessalegn Tadesse. 2015. Estimation of Heritability and Genetic Advance of Yield and Yield Related Traits in Bread Wheat (*Triticum aestivum* L.) Genotypes at Ofla District, Northern Ethiopia. *International Journal of Plant Breeding and Genetics* 10 (1): 31–37. https://doi.org/10.3923/ijpbg.2016.31.37.

Meseret, Asmamaw, Keneni Gemechu, and Tesfaye Kassahun. 2020. Genetic Diversity of Ethiopian Durum Wheat (*Triticum durum* Desf.) Landrace Collections as Reveled by Morphological Markers." *Journal of Plant Breeding and Crop Science* 12 (04): 258–68. https://doi.org/10.4172/2329-8863.1000413.

Meuwissen, T H E, B J Hayes, and M E Goddard. 2001. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157 (4): 1819—1829.

Mir Drikvand, Reza, Goodarz Najafian, Mohammad Reza Bihamta, and Asa Ebrahimi. 2018. "Mapping Some Seed Quality Traits in Bread Wheat (*Triticum aestivum* L.) by Association Mapping Using Ssr Markers. *Journal of Applied Biotechnology Reports* 5 (3): 92–99. https://doi.org/10.29252/JABR.05.03.02.

Modoran, Constanta, Simona Man, V Mure, and Anamaria Birou. 2010. The Influence of Wheat Genotype on Its Quality 16 (June 2015): 99–103.

Morris C Peter, Bryce H James. 2000. Carbohydrate Polymers. In *Cereal Biotechnology*, 46:396–97. https://doi.org/10.1016/s0144-8617(01)00254-5.

Muleta, Kebede T, Matthew N Rouse, Sheri Rynearson, Xianming Chen, Bedada G Buta, and Michael O Pumphrey. 2017. Characterization of Molecular Diversity and Genome-Wide Mapping of Loci Associated with Resistance to Stripe Rust and Stem Rust in Ethiopian Bread Wheat Accessions. *BMC Plant Biology* 17 (134): 1–20. https://doi.org/10.1186/s12870-017-1082-7.

Mulugeta, B. *et al.* (2023) 'Harnessing genome-wide genetic diversity, population structure and linkage disequilibrium in Ethiopian durum wheat gene pool', *Frontiers in Plant Science*, 14(July), pp. 1–18. Available at: https://doi.org/10.3389/fpls.2023.1192356.

Mwadzingeni, Learnmore, Hussein Shimelis, D Jasper G Rees, and Toi J Tsilo. 2017. "Genome-Wide Association Analysis of Agronomic Traits in Wheat under Drought-Stressed and Non-Stressed Conditions. *PLoS ONE* 12 (2): 1–13.

https://doi.org/10.1371/journal.pone.0171692.
N.Wang, F. E. Dowell, and N. Zhang. 2002. Determining Wheat Vitreousness Using Image Processing and a Neural Network. *Transactions of the American Society of Agricultural Engineers* 46 (4): 1143–50. https://doi.org/10.13031/2013.13937.
Nakaya, Akihiro, and Sachiko N Isobe. 2012. Will Genomic Selection Be a Practical Method for Plant Breeding ? *Annals of Botany* 110: 1303–16. https://doi.org/10.1093/aob/mcs109.
Nalam, Vamsi J, Æ M Isabel Vales, Christy J W Watson, and Æ Shahryar F Kianian. 2006. "Map-Based Analysis of Genes Affecting the Brittle Rachis Character in Tetraploid Wheat (*Triticum turgidum* L .), 373–81. https://doi.org/10.1007/s00122-005-0140-y.
Ne, Masatoshi. 1978. Estimation of Average Heterozygosity and Genetic Distance from a Small Number of Individuals" *Genetics*, 583–90.
Neethirajan, S., C. Karunakaran, S. Symons, and D. S. Jayas. 2006. Classification of Vitreousness in Durum Wheat Using Soft X-Rays and Transmitted Light Images. *Computers and Electronics in Agriculture* 53 (1): 71–78. https://doi.org/10.1016/j.compag.2006.03.001.
Negash Geleta. 2015. Patterns of Variation for Phenotypic Traits in Tetraploid Wheat (*Triticum turgidum* L.) Populations of Ethiopia. *Agricultural and Biological Sciences Journal* 1 (2): 42–51.
Negash, Tamrat, Erena Aka Edae, Lidiya Tilahun, Prabin Bajgain, Matthew N Rouse, and James A Anderson. 2022. Genome-Wide Association Mapping for Field and Seedling Resistance to the Emerging Puccinia Graminis f . Sp . Tritici Race TTRTF in Wheat. *Plant Genome*, 1–14. https://doi.org/10.1002/tpg2.20274.
Negassa, M. 1986. Patterns of Diversity of Ethiopian Wheats (*Triticum spp*.) and a Gene Center for Quality Breeding. *Plant Breeding* 97 (2): 147–62. https://doi.org/10.1111/j.1439-0523.1986.tb01046.x.
Negessa, Mulugeta. 1986. Estimates of Phenotypic Diversity and Breeding Potential of Ethiopian Wheats. *Hereditas* 104 (1): 41–48. https://doi.org/10.1111/j.1601-5223.1986.tb00515.x.
Negisho, Kefyalew, Surafel Shibru, Klaus Pillenid, Frank and Ordon, and Gwendolin Wehnerid. 2021. Genetic Diversity of Ethiopian Durum Wheat Landraces. *PLoS ONE* 16 (2): 1–15. https://doi.org/10.1371/journal.pone.0247016.
Nei, M. 1977. F-Statistics and Gene Diversity in Subdivided Populations. *Ann. Hum. Gent.* 41:221-233: 225–33.
Nigus, Mulu, Hussein Shimelis, Isack Mathew, and Seltene Abady. 2022. Wheat Production in the Highlands of Eastern Ethiopia: Opportunities, Challenges and Coping Strategies of Rust Diseases. *Acta Agriculturae Scandinavica Section B: Soil and Plant Science* 72. https://doi.org/10.1080/09064710.2021.2022186.
Oliveira, Hugo R, Michael G Campana, Huw Jones, Harriet V Hunt, Fiona Leigh, David I Redhouse, Diane L Lister, and Martin K Jones. 2012. Tetraploid Wheat Landraces in the Mediterranean Basin: Taxonomy, Evolution and Genetic Diversity. *PLoS ONE* 7 (5): 1–13. https://doi.org/10.1371/journal.pone.0037063.
Oraguze, N.C., P.L. Wilcox, E.H.A. Rikkerink, and H.N. de Silva. 2007. *Association Mapping in Plants*. Edited by C.N. Oraguze, E.H.A. Rikkerink, S.E. Gardiner, and H.N De Silva. *North*. New York: Springer Sciences Business Media LLC.
Özkan, Hakan, George Willcox, Andreas Graner, Francesco Salamini, and Benjamin Kilian. 2011. Geographic Distribution and Domestication of Wild Emmer Wheat (*Triticum dicoccoides*). *Genetic Resources and Crop Evolution* 58 (1): 11–53. https://doi.org/10.1007/s10722-010-9581-5.
Pasam, Raj K, Rajiv Sharma, Marcos Malosetti, Fred A Van Eeuwijk, Grit Haseneyer,

Benjamin Kilian, and Andreas Graner. 2012. Genome-Wide Association Studies for Agronomical Traits in a World Wide Spring Barley Collection. *BMC Plant Biology* 12 (16): 1–22. https://doi.org/10.1186/1471-2229-12-16.

Peakall, Rod, and Peter E. Smouse. 2012. GenALEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research-an Update. *Bioinformatics* 28 (19): 2537–39. https://doi.org/10.1093/bioinformatics/bts460.

Pecetti, L, A B Damania, and G Kashour. 1992. Geographic Variation for Spike and Grain Characteristics in Durum Wheat Germplasm Adapted to Dryland Conditions, 97–105.

Pecetti, Luciano, and Ardeshir B. Damania. 1996. Geographic Variation in Tetraploid Wheat (*Triticum turgidum* Ssp. Turgidum Convar. Durum) Landraces from Two Provinces in Ethiopia. *Genetic Resources and Crop Evolution* 43 (5): 395–407.

Peña, R. J., R. Trethowan, W. H. Pfeiffer, and M. Van Ginkel. 2002. *Quality (End-Use) Improvement in Wheat*. *Journal of Crop Production*. Vol. 5. https://doi.org/10.1300/j144v05n01_02.

Peng, Junhua H, Dongfa Sun, and Eviatar Nevo. 2011. *Domestication Evolution, Genetics and Genomics in Wheat*. *Molecular Breeding*. Vol. 28. https://doi.org/10.1007/s11032-011-9608-4.

Persaud, Mahendra, Rajendra Persaud, Nandram Gobind, Ayub Khan, Gomathinayagam Subramanian, and Edgar Corredor. 2022. Genotype by Environment Interactions of Grain Yield Performance and Lodging Incidence in Advance Breeding Lines of Rice across Environments in Guyana. *International Journal of Agricultural Policy and Research* 10 (3): 70–81. https://doi.org/10.15739/ijapr.22.009.

Petrova, I. 2007. End-Use Quality of Bulgarian Durum Wheat. *Bulgarian Journal of Agricultural Science* 13 (2): 161–69.

Philips, R L, and I K Vasil. 2001. *DNA-Based Markers in Plants, 2nd Edition*. Vol. 1. http://www.springerlink.com/index/10.1007/978-94-011-1104-1.

Prasad, J., A. Dasora, D. Chauhan, D.A. Rizzardi, Bangarwa S. K., and K. Nesara. 2023. Genetic Variability, Heritability and Genetic Advance Analysis in Bread Wheat (*Triticum aestivum* L.) Genotypes. *International Journal of Plant & Soil Science* 35 (19): 164–72. https://doi.org/10.9734/ijpss/2023/v35i193538.

R. Harlan. 1971. Agricultural Origins : Centers and Noncenters. *Science*.

R Development Core Team, R. 2018. R: A Language and Environment for Statistical Computing. Edited by R Development Core Team. *R Foundation for Statistical Computing*. R Foundation for Statistical Computing. R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7.

Rafalski, Antoni. 2002. Applications of Single Nucleotide Polymorphisms in Crop Genetics. *Current Opinion in Plant Biology* 5 (2): 94–100. https://doi.org/10.1016/s1369-5266(02)00240-6.

Rahimi, Yousef, Mohammad Reza Bihamta, Alireza Taleei, Hadi Alipour, and Pär K. Ingvarsson. 2019. Genome-Wide Association Study of Agronomic Traits in Bread Wheat Reveals Novel Putative Alleles for Future Breeding Programs. *BMC Plant Biology* 19 (1): 1–19. https://doi.org/10.1186/s12870-019-2165-4.

Ramsay, Luke, Jordi Comadran, Arnis Druka, David F. Marshall, William T.B. Thomas, Malcolm MacAulay, Katrin MacKenzie, et al. 2011. INTERMEDIUM-C, a Modifier of Lateral Spikelet Fertility in Barley, Is an Ortholog of the Maize Domestication Gene TEOSINTE BRANCHED 1. *Nature Genetics* 43 (2): 169–72. https://doi.org/10.1038/ng.745.

Remington, David L, Jeffry M Thornsberry, Yoshihiro Matsuoka, Larissa M Wilson, Sherry R Whitt, John Doebley, Stephen Kresovich, Major M Goodman, and Edward S Buckler Iv. 2001. Structure of Linkage Disequilibrium and Phenotypic Associations

in the Maize Genome. *Proceedings of the National Academy of Sciences of the United States of America* 98 (20): 11479–84. https://doi.org/10.1073/pnas.201394398.

Ren, Jing, Daokun Sun, Liang Chen, Frank M You, Jirui Wang, and Yunliang Peng. 2013. Genetic Diversity Revealed by Single Nucleotide Polymorphism Markers in a Worldwide Germplasm Collection of Durum Wheat. *International Journal of Molecular Sciences* 14 (February): 7061–88. https://doi.org/10.3390/ijms14047061.

Robbana, Cyrine, Zakaria Kehel, M'Barek Ben Naceur, Carolina Sansaloni, Filippo Bassi, and Ahmed Amri. 2019. Genome-Wide Genetic Diversity and Population Structure of Tunisian Durum Wheat Landraces Based on Dartseq Technology. *International Journal of Molecular Sciences* 20 (6): 1–21. https://doi.org/10.3390/ijms20061352.

Roncallo, Pablo F, Pavan C Akkiraju, Gerardo L Cervigni, and Viviana C Echenique. 2017. QTL Mapping and Analysis of Epistatic Interactions for Grain Yield and Yield-Related Traits in *Triticum turgidum* L . Var . Durum. *Euphytica* 213 (277): 1–20. https://doi.org/10.1007/s10681-017-2058-2.

Ross-Ibarra, Jeffrey, Peter L. Morrell, and Brandon S. Gaut. 2007. Plant Domestication, a Unique Opportunity to Identify the Genetic Basis of Adaptation. *In the Light of Evolution* 1: 205–23. https://doi.org/10.17226/11790.

Rostoks, Nils, Luke Ramsay, Katrin MacKenzie, Linda Cardle, Prasanna R. Bhat, Mikeal L. Roose, Jan T. Svensson, et al. 2006. Recent History of Artificial Outcrossing Facilitates Whole-Genome Association Mapping in Elite Inbred Crop Varietes. *Proceedings of the National Academy of Sciences of the United States of America* 103 (49): 18656–61. https://doi.org/10.1073/pnas.0606133103.

Royo, Conxita, Miloudi M.Nachit, Natale Di Fonzo, Jose Luis Araus, Wolfgang H. Pfeiffer, and Gustavo A. Slafer. 2005. *Durum Wheat Breeding: Current Approaches and Future Strategies*. Edited by Conxita Royo, M.M. Nachit, N.Di Fonzo, J.L. Araus, W.H. Pfeiffer, and Gustavo A. Slafer. *Durum Wheat Breeding*. New York: Food Products Press.

Ruiz, Magdalena, Patricia Giraldo, Conxita Royo, Dolors Villegas, M. Jose Aranzana, and Jose M. Carrillo. 2012. Diversity and Genetic Structure of a Collection of Spanish Durum Wheat Landraces. *Crop Science* 52 (5): 2262–75. https://doi.org/10.2135/cropsci2012.02.0081.

Rushton, Paul J., Imre E. Somssich, Patricia Ringler, and Qingxi J. Shen. 2010. WRKY Transcription Factors. *Trends in Plant Science* 15 (5): 247–58. https://doi.org/10.1016/j.tplants.2010.02.006.

Sajjad, Muhammad, Sultan Habibullah Khan, and Munawar Shahzad. 2018. Patterns of Allelic Diversity in Spring Wheat Populations by SSR-Markers. *Cytology and Genetics* 52 (2): 155–60. https://doi.org/10.3103/S0095452718020081.

Salamini, Francesco, Hakan Özkan, Andrea Brandolini, Ralf Schäfer-Pregl, and William Martin. 2002. Genetics and Geography of Wild Cereal Domestication in the near East. *Nature Reviews Genetics* 3 (6): 429–41. https://doi.org/10.1038/nrg817.

Salem, Khaled F.M., and Ahmed Sallam. 2016. Analysis of Population Structure and Genetic Diversity of Egyptian and Exotic Rice (*Oryza sativa* L.) Genotypes. *Comptes Rendus - Biologies* 339 (1): 1–9. https://doi.org/10.1016/j.crvi.2015.11.003.

Salvador-Figueroa, M, J Magaña-Ramos, J A Vázquez-Ovando, M L Adriano-Anaya, and I Ovando-Medina. 2015. Genetic Diversity and Structure of Jatropha Curcas L. in Its Centre of Origin. *Plant Genetic Resources: Characterisation and Utilisation* 13 (1): 1–21. https://doi.org/10.1017/S1479262114000550.

Sara, Farokhzadeh, Ali Fakheri Barat, Nafiseh Mahdi Nezhad, Tahmasebi Sirous, Sara Farokhzadeh, Sirous Tahmasebi Barat Ali FakheriA, Nafiseh Mahdi NezhadA, Mirsoleimani Abbas, and Bahram Heidari. 2020. Mapping QTLs Associated with

Grain Yield and Yield-Related Traits under Aluminum Stress in Bread Wheat. *Crop and Pasture Science* 71: 429–44. https://doi.org/https://doi.org/10.1071/CP19511.

Sayaslan, A. 2012. Some Quality Characteristics of Selected Durum Wheat (*Triticum durum*) Landraces. *Turkish Journal of Agriculture and Forestry* 36 (749–756). https://doi.org/10.3906/tar-1201-69.

Sharma, Subarna, Nav R. Acharya, Sharad Adhikari, &, and Krishna K. Mishra. 2017. Varietal Improvement of Wheat Under Rainfed Conditions in Mid-Western Terai of Nepal. *Global Journal of Biology, Agriculture & Health Sciences* 6 (4): 15–19. https://doi.org/10.24105/gjbahs.6.4.1704.

Shewry, P R. 2009. Wheat. *Journal of Experimental Botany* 60 (6): 1537–53. https://doi.org/10.1093/jxb/erp058.

Simons, Kristin J., John P. Fellers, Harold N. Trick, Zengcui Zhang, Yin Shan Tai, Bikram S. Gill, and Justin D. Faris. 2006. Molecular Characterization of the Major Wheat Domestication Gene Q. *Genetics* 172 (1): 547–55. https://doi.org/10.1534/genetics.105.044727.

Singh, R. K., and B. D. Chaudhary. 1985. *Biometrical Methods in Quantitative Genetic Analysis.* Kalyani. https://doi.org/10.2307/2530404.

Sneath, P.H. and Sokal, R.R. 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classification.

Somers, D.J., Peter Isaac, and Keith Edwards. 2004. A High-Density Microsatellite Consensus Map for Bread Wheat (*Triticum aestivum* L .). *Theoretical and Applied Genetics* 109: 1105–14. https://doi.org/10.1007/s00122-004-1740-7.

Stich, Benjamin, Hans P. Maurer, Albrecht E. Melchinger, Matthias Frisch, Martin Heckenberger, Jeroen Rouppe Van Der Voort, Johan Peleman, Anker P. Sørensen, and Jochen C. Reif. 2006. Comparison of Linkage Disequilibrium in Elite European Maize Inbred Lines Using AFLP and SSR Markers. *Molecular Breeding* 17 (3): 217–26. https://doi.org/10.1007/s11032-005-5296-2.

Studnicki, Marcin, Magdalena Wijata, Grzegorz Sobczyński, Stanisław Samborski, Dariusz Gozdowski, and Jan Rozbicki. 2016. Effect of Genotype, Environment and Crop Management on Yield and Quality Traits in Spring Wheat. *Journal of Cereal Science* 72: 30–37. https://doi.org/10.1016/j.jcs.2016.09.012.

Sukumaran, Sivakumar, Susanne Dreisigacker, Marta Lopes, Perla Chavez, and Matthew P Reynolds. 2014. Genome - Wide Association Study for Grain Yield and Related Traits in an Elite Spring Wheat Population Grown in Temperate Irrigated Environments. *Theoretical and Applied Genetics*. https://doi.org/10.1007/s00122-014-2435-3.

Sun, Xiaoxiao, Tongtong Liu, Tangyuan Ning, Kai Liu, Xixian Duan, Xuerong Wang, Qiaoling Wang, et al. 2018. Genetic Dissection of Wheat Kernel Hardness Using Conditional QTL Mapping of Kernel Size and Protein-Related Traits. *Plant Molecular Biology Reporter* 36: 1–12. https://doi.org/10.1007/s11105-017-1061-z.

Talini, Rebecca F., Andrea Brandolini, Mara Miculan, Alice Brunazzi, Patrizia Vaccino, Mario Enrico Pe, and Matteo Dell'Acqua. 2020. Genome-Wide Association Study of Agronomic and Quality Traits in a World Collection of the Wild Wheat Relative *Triticum urartu*. *Plant Journal* 102: 555–68. https://doi.org/10.1111/tpj.14650.

Taneva, Krasimira, Violeta Bozhanova, and Ivanka Petrova. 2019. Variability, Heritability and Genetic Advance of Some Grain Quality Traits and Grain Yield in Durum Wheat Genotypes. *Bulgarian Journal of Agricultural Science* 25 (2): 288–95.

Tank, J. G., and V. S. Thaker. 2011. Cyclin Dependent Kinases and Their Role in Regulation of Plant Cell Cycle. *Biologia Plantarum* 55 (2): 201–12. https://doi.org/10.1007/s10535-011-0031-9.

Teklu, Yifru, K. Hammer, X. Q. Huang, and M .S. Roder. 2006. Analysis of Microsatellite Diversity in Ethiopian Tetraploid Wheat Landraces. *Genetic Resources and Crop Evolution* 53: 1115–26. https://doi.org/10.1007/s10722-005-1146-7.

Teklu, Yifru, and Karl Hammer. 2008. Diversity of Ethiopian Tetraploid Wheat Germplasm: Breeding Opportunities for Improving Grain Yield Potential and Quality Traits. *Plant Genetic Resources: Characterisation and Utilisation* 7 (1): 1–8. https://doi.org/10.1017/S1479262108994223.

Tesemma, Tesfaye, H C Becker, Belay G, Mitiku D, E Bechere, and S Tsegaye. 1993. Performance of Ethiopian Tetraploid Wheat Landraces at Their Collection Sites. *Euphytica* 71: 221–30.

Thapa, R., G. Brown-Guedira, H.W. Ohm, K. Wise, and S.B Goodwin. 2015. *Advances in Wheat Genetics: From Genome to Field*. Edited by Y. Ogihara, S. Takumi, and H. Handa. *Determining the Order of Resistance Genes Qsng-3BS, Fhb1 and Sr2 and Combining Them in Coupling on Wheat Chromosome 3BS*. Springer Sciences Business Media LLC. https://doi.org/10.1007/978-4-431-55675-6_36.

Thornsberry, Jeffry M., Major M. Goodman, John Doebley, Stephen Kresovich, Dahlia Nielsen, and Edward S. Buckler. 2001. Dwarf8 Polymorphisms Associate with Variation in Flowering Time. *Nature Genetics* 28 (3): 286–89. https://doi.org/10.1038/90135.

Tian, Jichun, Jiansheng Chen, Guangfeng Chen, Peng Wu, Han Zhang, and Yong Zhao. 2015. *Genetic Analyses of Wheat and Molecular Marker-Assisted Breeding, Volume 2*. https://doi.org/10.1007/978-94-017-7447-5.

Tuberosa, Roberto. 2012. Phenotyping for Drought Tolerance of Crops in the Genomics Era. *Fronteirs in PHYSIOLOGY* 3 (Article 347): 1–27. https://doi.org/10.3389/fphys.2012.00347.

Turki, Nejla, Tariq Shehzad, Moncef Harrabi, and KAzutoshi Okuno. 2014. Detection of QTLs Associated with Salinity Tolerance in Durum Wheat Based on Association Analysis. *Euphytica*. https://doi.org/10.1007/s10681-014-1164-7.

Tzarfati, R, V Barak, T Krugman, T Fahima, S Abbo, Y Saranga, and A B Korol. 2014. Novel Quantitative Trait Loci Underlying Major Domestication Traits in Tetraploid Wheat. *Molecular Breeding* 34: 1613–28. https://doi.org/10.1007/s11032-014-0182-4.

Vanraden, P M. 2008. Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* 91 (11): 4414–23. https://doi.org/10.3168/jds.2007-0980.

Varzakas, Theo, Natalia Kozub, and Ioannis N. Xynias. 2014. *Quality Determination of Wheat: Genetic Determination, Biochemical Markers, Seed Storage Proteins - Bread and Durum Wheat Germplasm*. *Journal of the Science of Food and Agriculture*. Vol. 94. https://doi.org/10.1002/jsfa.6601.

VILOV,N.I. 1950. *The Origin, Variation, Immunity and Breeding of Cultivated Plants. (1949, Chronica Botanica) - Libgen.Li.Pdf* 13 (1).

Vavilov, N.I. 1929. Origin and Geography of Cultivated Plants by N. I. Vavilov, Doris Love (z-Lib.Org).

Wang, Daowen, Kunpu Zhang, Lingli Dong, Zhenying Dong, Yiwen Li, Abrar Hussain, and Huijie Zhai. 2017. Molecular Genetic and Genomic Analysis of Wheat Milling and End-Use Traits in China: Progress and Perspectives. *Crop Journal* 6 (1): 68–81. https://doi.org/10.1016/j.cj.2017.10.001.

Wang, Han-Yan, Yu-Ming Wei, Ze-Hong Yan, and You-Liang Zheng. 2007. EST-SSR DNA Polymorphism in Durum Wheat. *Journal of Applied Genetics* 48 (1): 35–42. https://doi.org/10.1007/bf03194655.

Wang, Jun, and Bingkai Hou. 2016. Glycosyltransferases : Key Players Involved in the Modi Fi Cation of Plant Secondary Metabolites. *Frontiers of Biology in China* 4: 39–

46. https://doi.org/10.1007/s11515-008-0111-1.
Wang, Meng, Jiarui Yuan, Lumin Qin, Weiming Shi, Guangmin Xia, and Shuwei Liu. 2020. TaCYP81D5, One Member in a Wheat Cytochrome P450 Gene Cluster, Confers Salinity Tolerance via Reactive Oxygen Species Scavenging. *Plant Biotechnology Journal* 18 (3): 791–804. https://doi.org/10.1111/pbi.13247.
Wang, N, F E Dowell, and N Zhang. 2002. Determining Wheat Vitreousness Using Image Processing and a Neural Network. *Transactions of the American Society of Agricultural Engineers* 46 (4): 1143–50. https://doi.org/10.13031/2013.13937.
Wang, Shubin, Steven Xu, Shiaoman Chao, Qun Sun, Shuwei Liu, and Guangmin Xia. 2019. A Genome-Wide Association Study of Highly Heritable Agronomic Traits in Durum Wheat. *Frontiers in Plant Science* 10 (July): 1–13. https://doi.org/10.3389/fpls.2019.00919.
Wang, Xin, Guangbin Luo, Wenlong Yang, Yiwen Li, Jiazhu Sun, Kehui Zhan, Dongcheng Liu, and Aimin Zhang. 2017. Genetic Diversity, Population Structure and Marker-Trait Associations for Agronomic and Grain Traits in Wild Diploid Wheat *Triticum urartu*. *BMC Plant Biology* 17 (1): 1–17. https://doi.org/10.1186/s12870-017-1058-7.
Wang, Yujiao, Yong Zhang, Chunjie Fan, Yongcheng Wei, Jingxiang Meng, Zhen Li, and Chonglu Zhong. 2021. Genome-Wide Analysis of MYB Transcription Factors and Their Responses to Salt Stress in Casuarina Equisetifolia. *BMC Plant Biology* 21 (1): 1–17. https://doi.org/10.1186/s12870-021-03083-6.
Wang, Zhilai, Yanfeng Zhang, Min Song, Xiuhua Tang, Shuhua Huang, Bin Linhu, Ping Jin, et al. 2023. Genome-Wide Identification of the Cytochrome P450 Superfamily Genes and Targeted Editing of BnCYP704B1 Confers Male Sterility in Rapeseed. *Plants* 12 (2). https://doi.org/10.3390/plants12020365.
Wenzl, Peter, Jason Carling, David Kudrna, Damian Jaccoud, Eric Huttner, Andris Kleinhofs, Andrzej Kilian, et al. 2016. Diversity Arrays Technology (DArT) for Whole-Genome Profiling of Barley. In *Proceedings of the National Academy of Sciences*, 101:9915–20. https://doi.org/10.1073/pnas.0401076101.
Werck-Reichhart D, Hehn A, and and Didierjean L. 2000. Cytochromes P450 for Engineering Herbicide Tolerance. *Trends in Plant Science* 5 (3): 116–23.
Werner, Jonathan D., Justin O. Borevitz, N. Henriette Uhlenhaut, Joseph R. Ecker, Joanne Chory, and Detlef Weigel. 2005. FRIGIDA-Independent Variation in Flowering Time of Natural Arabidopsis Thaliana Accessions. *Genetics* 170 (3): 1197–1207. https://doi.org/10.1534/genetics.104.036533.
Wright, Sewall. 1978. *Sewall Wright - Evolution and Genetics of Populations*.
Xu, Jun, Xin Yu Wang, and Wang Zhen Guo. 2015. The Cytochrome P450 Superfamily: Key Players in Plant Development and Defense. *Journal of Integrative Agriculture* 14 (9): 1673–86. https://doi.org/10.1016/S2095-3119(14)60980-1.
Xu, Yunbi, and Jonathan H. Crouch. 2008.Marker-Assisted Selection in Plant Breeding: From Publications to Practice. *Crop Science* 48 (2): 391–407. https://doi.org/10.2135/cropsci2007.04.0191.
Yalcin, Seda, and Besim Maden. 2020. Comparison of Protein Characteristics of Some Wheat Varieties Harvested in Turkey. *Afyon Kocatepe University Journal of Sciences and Engineering* 20 (3): 441–47. https://doi.org/10.35414/akufemubid.606621.
Yao, Yanjuan, Qinhuo Liu, Qiang Liu, and Xiaowen Li. 2008. LAI Retrieval and Uncertainty Evaluations for Typical Row-Planted Crops at Different Growth Stages. *Remote Sensing of Environment* 112: 94–106. https://doi.org/10.1016/j.rse.2006.09.037.
Yematawy, Z., G. Blomme, S. Muzemil, and K. Tesfaye. 2018. Assessing Qualitative and

Phenotypic Trait Diversity in Ethiopian Enset [*Ensete ventricosum* (Welw.) Cheesman] Landraces. *Fruits* 73 (6): 310–27. https://doi.org/10.17660/th2018/73.6.2.

Yue, Hong, Duntao Shu, Meng Wang, Guangwei Xing, Haoshuang Zhan, Xianghong Du, Weining Song, and Xiaojun Nie. 2018. Genome-Wide Identification and Expression Analysis of the Hd-Zip Gene Family in Wheat (*Triticum aestivum* L.). *Genes* 9 (2): 1–19. https://doi.org/10.3390/genes9020070.

Zegeye, Fisseha, Bamlaku Alamirew, and Degefa Tolossa. 2020. Analysis of Wheat Yield Gap and Variability in Ethiopia. *International Journal of Agricultural Economics* 5 (4): 89–98. https://doi.org/10.11648/j.ijae.20200504.11.

Zewdie B., Prof. dr. ir. P C Struik, Dr. ir. A J G van Gastel, and Z Bishaw. 2004. *Wheat and Barley Seed Systems in Ethiopia and Syria*. *Production Ecology and Resource Conservation*. Vol. Phd Thesis. https://library.wur.nl/WebQuery/wurpubs/fulltext/121548.

Zhang, Liyi, Dongcheng Liu, Xiaoli Guo, Wenlong Yang, Jiazhu Sun, Daowen Wang, and Pierre Sourdille. 2011. Investigation of Genetic Diversity and Population Structure of Common Wheat Cultivars in Northern China Using DArT Markers. *BMC Evolutionary Biology* 12: 1–11. https://doi.org/10.1186/1471-2156-12-42.

Zhang, Shulin, Zailong Tian, Haipeng Li, Yutao Guo, Yanqi Zhang, Jeremy A Roberts, Xuebin Zhang, and Yuchen Miao. 2019. Genome-Wide Analysis and Characterization of F-Box Gene Family in *Gossypium hirsutum* L" *BMC Genetics*, 1–16.

Zhang, W., L. J. Qu, H. Gu, W. Gao, M. Liu, J. Chen, and Z. Chen. 2002. Studies on the Origin and Evolution of Tetraploid Wheats Based on the Internal Transcribed Spacer (ITS) Sequences of Nuclear Ribosomal DNA. *Theoretical and Applied Genetics* 104 (6–7): 1099–1106. https://doi.org/10.1007/s00122-002-0887-3.

Zheng, Peizhong, William B. Allen, Keith Roesler, Mark E. Williams, Shirong Zhang, Jiming Li, Kimberly Glassman, et al. 2008. A Phenylalanine in DGAT Is a Key Determinant of Oil Content and Composition in Maize. *Nature Genetics* 40 (3): 367–72. https://doi.org/10.1038/ng.85.

Zhu, Chengsong, Michael Gore, Edward S. Buckler, and Jianming Yu. 2008. Status and Prospects of Association Mapping in Plants. *The Plant Genome* 1 (1): 5–20. https://doi.org/10.3835/plantgenome2008.02.0089.

Zohary, Daniel, and Moshe Feldman. 1962. Hybridization Between Amphidiploids and the Evolution of Polyploids in the Wheat (*Aegilops-triticum*) Group. *Evolution* 16 (1): 44. https://doi.org/10.2307/2406265.

Appendixes

Annex 1 seed source, pedigree, and passport data of tetraploid wheat of Ethiopia germplasm used in the genetic diversity study

Genotype	Germplasm	**Region**	**Zone**	**Altitude**	**Source**	Genotype	Germplasm	**Region**	**Zone**	**Altitude**	**Source**
ETW1	226694			2150	SARC	ETW66	5576	Oromiya	North SHEWA	2510	SARC
ETW2	5760	Oromiya	JIMMA	2667	SARC	ETW67	5600	Oromiya	North SHEWA	3050	SARC
ETW3	6987	Oromiya	ARSSI	2580	SARC	ETW68	226241	Oromiya	North SHEWA	2900	SARC
ETW4	222427	Oromiya	ARSSI	2460	SARC	ETW69	5584	Oromiya	North SHEWA	2650	SARC
ETW5	7762	Oromiya	ARSSI	3000	SARC	ETW70	5581	Oromiya	North SHEWA	2520	SARC
ETW6	7004	Oromiya	ARSSI	1660	SARC	ETW71	DW17007	Amara	Wello	2678	DZARC
ETW7	7050	Oromiya	ARSSI	2600	SARC	ETW72	5369	Oromiya	North SHEWA	2510	SARC
ETW8	7076	Oromiya	ARSSI	2430	SARC	ETW73	5638	Oromiya	North SHEWA	2160	SARC
ETW9	243701	Amara	N/ WELLO	2950	SARC	ETW74	5586	Oromiya	North SHEWA	2810	SARC
ETW10	7003	Oromiya	ARSSI	2760	SARC	ETW75	5609	Oromiya	N/SHEWA	2510	SARC
ETW11	7242	Oromiya	ARSSI	2780	SARC	ETW76	7191	Oromiya	N/SHEWA	2500	SARC
ETW12	7060	Oromiya	ARSSI	2520	SARC	ETW77	7150	Oromiya	NSHEWA	2735	SARC
ETW13	7056	Oromiya	ARSSI	2500	SARC	ETW78	5510	Oromiya	N/SHEWA	2420	SARC
ETW14	7014	Oromiya	ARSSI	2490	SARC	ETW79	5142	Oromiya	N/SHEWA	2893	SARC
ETW15	242789	Oromiya	E/SHEWA	2330	SARC	ETW80	7195	Oromiya	N/SHEWA	2510	SARC
ETW16	7313	Oromiya	ARSSI	2570	SARC	ETW81	5618	Oromiya	N/SHEWA	2810	SARC
ETW17	7069	Oromiya	ARSSI	2470	SARC	ETW82	5707	Oromiya	North SHEWA	2810	SARC
ETW18	5057	Oromiya	ARSSI	2500	SARC	ETW83	7626	Oromiya	North SHEWA	2950	SARC
ETW19	7015	Oromiya	ARSSI	2395	SARC	ETW84	5171	Amara	N/SHEWA	2260	SARC
ETW20	6955	Oromiya	ARSSI	2440	SARC	ETW85	5169	Amara	North SHEWA	2260	SARC
ETW21	7063	Oromiya	ARSSI	2500	SARC	ETW86	5728	Amara	North SHEWA	3100	SARC
ETW22	7020	Oromiya	ARSSI	2430	SARC	ETW87	5168	Amara	N/SHEWA	2260	SARC
ETW23	7078	Oromiya	ARSSI	2440	SARC	ETW88	5893	Amara	N/SHEWA	2620	SARC
ETW24	7007	Oromiya	ARSSI	2635	SARC	ETW89	5172	Amara	N/SHEWA	2260	SARC
ETW25	7629	Oromiya	NSHEWA	2950	SARC	ETW90	7713	Amara	N/SHEWA	2800	SARC
ETW26	7046	Oromiya	ARSSI	2620	SARC	ETW91	5166	Amara	N/SHEWA	2260	SARC
ETW27	242792	Oromiya	E/SHEWA	2330	SARC	ETW92	5184	Oromiya	E/SHEWA	2300	SARC
ETW28	7009	Oromiya	ARSSI	2150	SARC	ETW93	5197	Amara	AGEW AWI	2489	SARC
ETW29	239711	Oromiya	BALE	2760	SARC	ETW94	7825	Amara	East GOJAM	2460	SARC

ETW30	6933	Oromiya	BALE	2580	SARC	ETW95	7571	Amara	S/WELLO	2530	SARC
ETW31	238891	Oromiya	BALE	1900	SARC	ETW96	7568	Amara	S/WELLO	2680	SARC
ETW32	5149	Oromiya	W/SHEWA	2330	SARC	ETW97	243706	Amara	South WELLO	2825	SARC
ETW33	6102	Oromiya	W/SHEWA	2430	SARC	ETW98	7532	Amara	South WELLO	2300	SARC
ETW34	5554	Oromiya	W/SHEWA	2333	SARC	ETW99	7578	Amara	South WELLO	2540	SARC
ETW35	7999	Oromiya	W/SHEWA	2260	SARC	ETW100	7378	Amara	South WELLO	2400	SARC
ETW36	7209	Oromiya	W/SHEWA	2500	SARC	ETW101	7581	Amara	South WELLO	2510	SARC
ETW37	5182	Oromiya	W/SHEWA	2300	SARC	ETW102	7464	Amara	South WELLO	2500	SARC
ETW38	5591	Oromiya	E/SHEWA	2300	SARC	ETW103	243700	Amara	South WELLO	2940	SARC
ETW39	242786	Oromiya	E/SHEWA	2345	SARC	ETW104	235051	Amara	South WELLO	3100	SARC
ETW40	5666	Oromiya	E/SHEWA	2080	SARC	ETW105	7507	Amara	N/WELLO	2850	SARC
ETW41	5534	Oromiya	E/SHEWA	2350	SARC	ETW106	7477	Amara	N/WELLO	2900	SARC
ETW42	5491	Oromiya	E/SHEWA	2300	SARC	ETW107	243703	Amara	North WELLO	2950	SARC
ETW43	5044	Oromiya	E/SHEWA	1773	SARC	ETW108	7826	Amara	E/GOJAM	2460	SARC
ETW44	242790	Oromiya	E/SHEWA	2330	SARC	ETW109	7832	Amara	East GOJAM	2440	SARC
ETW45	5465	Oromiya	E/SHEWA	2800	SARC	ETW110	7673	Amara	E/GOJAM	2560	SARC
ETW46	5729	Oromiya	E/SHEWA	2300	SARC	ETW111	6971	Amara	East GOJAM	2610	SARC
ETW47	5492	Oromiya	E/SHEWA	2300	SARC	ETW112	7641	Amara	East GOJAM	2519	SARC
ETW48	242784	Oromiya	E/SHEWA	2340	SARC	ETW113	7827	Amara	East GOJAM	2460	SARC
ETW49	242781	Oromiya	E/SHEWA	2120	SARC	ETW114	7664	Amara	East GOJAM	2430	SARC
ETW50	5913	Oromiya	E/SHEWA	1915	SARC	ETW115	7683	Amara	East GOJAM	2450	SARC
ETW51	5174	Oromiya	E/SHEWA	2300	SARC	ETW116	6975	Amara	E/GOJAM	2550	SARC
ETW52	5635	Oromiya	E/SHEWA	2300	SARC	ETW117	7798	Amara	E/GOJAM	2520	SARC
ETW53	7880	Oromiya	E/SHEWA	2400	SARC	ETW118	7647	Amara	E/GOJAM	2470	SARC
ETW54	5515	Oromiya	E/SHEWA	2300	SARC	ETW119	6968	Amara	E/GOJAM	2460	SARC
ETW55	5441	Oromiya	E/SHEWA	2300	SARC	ETW120	6982	Amara	E/GOJAM	2600	SARC
ETW56	214370	Oromiya	E/SHEWA	1975	SARC	ETW121	7666	Amara	E/GOJAM	2410	SARC
ETW57	5052	Oromiya	E/SHEWA	2680	SARC	ETW122	7690	Amara	West GOJAM	2840	SARC
ETW58	7205	Oromiya	E/SHEWA	2500	SARC	ETW123	5342	Amara	North GONDAR	2720	SARC
ETW59	5373	Oromiya	E/SHEWA	1791	SARC	ETW124	5470	Amara	N/GONDAR	3020	SARC
ETW60	242793	Oromiya	E/SHEWA	2330	SARC	ETW125	5216	Amara	North GONDAR	2100	SARC
ETW61	5181	Oromiya	E/SHEWA	2300	SARC	ETW126	5214	Amara	N/GONDAR	2100	SARC
ETW62	5183	Oromiya	E/SHEWA	2300	SARC	ETW127	7295	Amara	South ONDAR	2640	SARC
ETW63	5434	Oromiya	E/SHEWA	2450	SARC	ETW128	5526	Tigray	DEBUBAWI	2687	SARC
ETW64	242782	Oromiya	E/SHEWA	2120	SARC	ETW129	243717	Tigray	MEHAKELEGNAW	2620	SARC
ETW65	5140	Oromiya	N/SHEWA	3066	SARC	*	7823	Amara	E/GOJAM	2460	SARC

*	5182	Oromiya	W/SHEWA	2300	SARC	*	5504	Oromiya	E/SHEWA	2300	SARC	
*	242783	Oromiya	E/SHEWA	2120	SARC	*	242783	Oromiya	E/SHEWA	2120	SARC	
*	7075	Oromiya	ARSSI	2430	SARC							
ETW130	MCD11-7,MCD11-25,MCD13-5,MCD14-19									2619	DZARC	
ETW131	ETDW15DZ012,ETDW/15DZ070			Oromiya		West SHEWA				2488	DZARC	
ETW132	ETDW/15DZ043,DZDW17010.8			Oromiya		West SHEWA				2496	DZARC	
ETW133	ETDW/15DZ#12,ETDW/15DZ#21……			Oromiya		West SHEWA				2498	DZARC	
ETW134	MCD4-4			Oromiya		East SHEWA				2522	DZARC	
ETW135	MCD11-20,MCD8-2,ETDW/15DZ019,……			Oromiya		East SHEWA				2711	DZARC	
ETW136	MCD12-2,MCD12-38,MCD14-26,…….			Oromiya		East SHEWA				2623	DZARC	
ETW137	MCD4-32IR70			Oromiya		East SHEWA				2702	DZARC	
ETW138	ETDW/15DZ#50,ETDW/15DZ#52,……			Oromiya		West SHEWA				2461	DZARC	
ETW139	MCD5-18			Oromiya		East SHEWA				2632	DZARC	
ETW140	DZDW176,7,10,DZDW17005,DZDW17006,07			Amara		Wello				2486	DZARC	
ETW141	MCD14-29			Oromiya		East SHEWA				2656	DZARC	
ETW142	MCD8-44			Oromiya		East SHEWA				2626	DZARC	
ETW143	ETDW/15DZ#044			Oromiya		West SHEWA				2480	DZARC	
ETW144	MCD12-30			Oromiya		East SHEWA				2601	DZARC	
ETW145	MCD3-19,ETDW/15DZ#34,ETDW/15DZ#35			Oromiya		West SHEWA				2495	DZARC	
ETW146	MCD10-11			Oromiya		East SHEWA				2566	DZARC	
ETW147	MCD3-14			Oromiya		East SHEWA				2609	DZARC	
ETW148	MCD3-27			Oromiya		East SHEWA				2682	DZARC	
ETW149	MCD4-13,MCD8-6,MCD11-18,…….			Oromiya		East SHEWA				2671	DZARC	
ETW150	ETDW15DZ073,ETDW/15DZ009,…….			Oromiya		West SHEWA				2454	DZARC	
ETW151	MCD14-3,MCD11-28,MCD11-36,…….			Oromiya		East SHEWA				2670	DZARC	
ETW152	DZDW17004,DZDW1702,04			Amara		Wello				2623	DZARC	
ETW153	MCD3-15			Oromiya		East SHEWA				2612	DZARC	
ETW154	DZDW17002			Amara		North SHEWA				2620	DZARC	
ETW155	ETDW/15DZ055,MCD3-1			Oromiya		West SHEWA				2454	DZARC	
ETW156	ETDW/15DZ#			Oromiya		East SHEWA				2423	DZARC	
ETW157	MCD5-4,MCD5-25,MCD6-44,……			Oromiya		East SHEWA				2677	DZARC	
ETW158	MCD13-8,MCD14-29			Oromiya		East SHEWA				2481	DZARC	
ETW159	MCD10-3,MCD10-16,MCD10-44……			Oromiya		East SHEWA				2487	DZARC	
ETW160	ETDW/15DZ#39			Oromiya		West SHEWA				2398	DZARC	
ETW161	MCD4-32LR72			Oromiya		East SHEWA				2678	DZARC	
ETW162	ETDW/15DZ#4,ETDW/15DZ#20……..			Oromiya		North SHEWA				2454	DZARC	
ETW163	MCD7-35			Oromiya		East SHEWA				2625	DZARC	
ETW164	ETDW/15DZ036,ETDW/15DZ063			Oromiya		West SHEWA				2372	DZARC	

ETW165	ETDW/15DZ#21ETDW/15DZ#31	Oromiya	West SHEWA	2655	DZARC
ETW166	MCD5-35,ETDW/15DZ004	Oromiya	East SHEWA	2454	DZARC
*	ETDW/15DZ#43ETDW/15DZ#068,…..			2461	DZARC
*	ETDW/15DZ014			2597	DZARC
	variety name	pedigree		year of release	Source
ETW167	Selam	61130/Lds//G11's'/3/cit's'/4/Hora/3/Megrbce's'		2004	DZARC
ETW168	Malefia			2005	DZARC
ETW169	Bekelcha	GEDLIFA/GUEROU		2005	DZARC
ETW170	Tob66	RCBCI/LD-357//DUCK''S''/ YEL''S'',TOB-66		1996	DZARC
ETW171	Quamy	ADS//PGO/CANDEAL II /7/JO''S''/ CR''S''//GS''S''/SBA81/3/FG''S''/ 4/FG''S''/CR''S''/5/ FG''S''DOM''S''/6/ HUI''S''CD75533-A		1996	DZARC
ETW172	Ejersa	LABUND/NIGRIS 3//GAN CD98206		2005	DZARC
ETW173	Fetan			2018	DZARC
ETW174	Ginchi	BOOHAI/ULNV - DZ1050		2000	DZARC
ETW175	ld-357	LD-357//CI8155 No 58-40		1979	DZARC
ETW176	Mangudo	ICAJIHAN 22		2012	DZARC
ETW177	Yerer			2004	DZARC
ETW178	Arendato			1967	DZARC
ETW179	DZ2018			2018	DZARC
ETW180	Boahi	Coo''S''/CANDEALL II CD3062- BSOGR		1982	DZARC
ETW181	**Cocorit/71**	RAE/4*TC60//STW63/3/ AA''s''DZ27617-18-64-OM		1976	DZARC
ETW182	Asasa	CHO''S''/TARUS//YAV''S''/3/ Fg''S''/4/Fgs/cr''S''/5/ DZ2085		1997	DZARC
ETW183	Gerardo	VZ 466/61-130xGII''s'', CM9605		1976	DZARC
ETW184	Foka	CIT71 CANDEAL II CD3369		1983	DZARC
ETW185	Tesfaye			2017	DZARC
ETW186	Werer			2009	DZARC
ETW187	Alemtena	Icasyr1/3/Gcn//Sti/Mrb3		2017	DZARC
ETW188	Mukiye	STJ3//BCR/LKS4/3/TER-3		2012	DZARC

* Genotypes not included in the genotyping by Sequencing study (GBS)

Annex 2 The percentage frequency of four qualitative traits of 174 tetraploid wheat landraces along with the 22 improved varieties based on the regions of collections, the species level, and the altitudinal classes

		GC			SD				Awnd			SC		
		White	RtoBr	PtoBl	VLx	Lx	Int	D	VD	AwnL	Awltd	CA	W	R	P
Population level		0.62	0.31	0.07	0.02	0.1	0.4	0.32	0.18				0.13	0.57	0.3
Region	Shewa	0.66	0.27	0.07	0.01	0.13	0.30	0.3	0.26		0.1	0.9	0.13	0.59	0.28
	Wello	0.63	0.26	0.11	0.05	0.05	0.42	0.42	0.05		0.11	0.89	0.11	0.42	0.47
	Arsi	0.42	0.54	0.04		0.04	0.54	0.38	0.04		0.08	0.92	0.04	0.65	0.31
	Gojam	0.65	0.26	0.09	0.04	0.04	0.52	0.3	0.1		0.04	0.96	0.22	0.48	0.3
Altitudinal class	≤1840	0.67	0.33		0.33		0.33	0.33						0.67	0.33
	1841-2020	1					0.33		0.67			1	0.33	0.33	0.33
	2021-2200	0.67	0.22	0.11		0.11	0.44	0.33	0.11			1	0.11	0.78	0.11
	2201-2380	0.50	0.43	0.07		0.13	0.33	0.3	0.23			1	0.06	0.67	0.27
	2381-2560	0.65	0.27	0.08		0.1	0.34	0.4	0.16		0.13	0.87	0.15	0.53	0.32
	2561-2740	0.63	0.29	0.07	0.02	0.07	0.46	0.22	0.22		0.11	0.89	0.15	0.51	0.34
	2741-2920	0.67	0.27	0.06		0.2	0.40	0.2	0.2		0.1	0.90	0.13	0.47	0.4
	≥2921	0.55	0.36	0.09		0.09	0.36	0.55				1	0.09	0.73	0.18
Species	DU	0.64	0.29	0.07		0.07	0.43	0.29	0.21		0.07	0.93	0.07	0.79	0.14
	TE	0.64	0.29	0.07	0.07	0.43	0.29	0.21			0.07	0.93	0.07	0.79	0.14
	TR	0.60	0.34	0.06	0.03	0.1	0.43	0.32	0.12		0.08	0.92	0.1	0.59	0.31
	Werer	100							100			100	100		
	Arendato	100							100			100	100		
	Tob66	100							100			100			100
	Yerer	100							100			100			100
	Gerardo		100						100			100			100
	Alemtena	100							100			100	100		
	Tesfaye	100							100			100			100
	Quamy		100						100			100			100
	Ld-357		100						100			100	100		

Ejersa	100	100	100		100
Foka	100	100	100		100
Malefia	100	100	100	100	
Mangudo	100	100	100		100
Fetan	100	100	100		100
DZ2018	100	100	100		100
Mukiye	100	100	100		100
Selam	100	100	100		100
Asasa	100	100	100		100
Bekelcha	100	100	100		100
Ginchi	100	100	100		100

NB: VL= very lax, L=Lax, int=intermediate, D=dense and VD= very dense; Awndness(AwL=awnless, Awltd=awnletted and CA=conspicuous awns); Seed colour (W=white, R=red and P=purple) and Glume colour(W=white, RtoB=red to brown and P=purple).

Annex 3 Mean values of phenotypic and grain quality traits of 196 genotypes evaluated at Sinana and Debrezeit during 2020

Phenotypic and grain quality trraits																		
Genotypes	PH	TPP	SpSp	SDpSp	SPL	TKW	GY	HI	BY	DB	DM	GC	SD	SC	Aw	GC	PC	ZI	VITP
5158	98.65	5.05	20.74	1.9	8.21	31.15	2.26	14.15	13.13	77	123.75	1	5	2	7	25.65	13.47	35.51	83.52
DZLR21	101	7.5	19.5	1.8	8.45	38.15	3.83	33.03	11.88	73.75	123	1	5	2	7	28.18	14.27	39.94	70.12
5526	101.45	5.05	19.2	2.2	9.11	30.75	2.3	20.42	10	79.25	127.5	2	7	3	7	25.94	12.97	38.44	61.42
7205	100.05	6.1	18.85	2	8.66	31.4	3.39	33.35	9.375	75.5	124.75	1	5	2	7	25.72	14.26	38.78	47.75
226241	104.6	5.4	18.45	2.3	8.29	35.95	3.77	36.73	10.63	74.75	126.75	3	7	3	7	25.84	12.42	37.49	71.35
7046	102.9	6.9	20.45	2.1	8.9	34.65	3.81	56.44	10.63	77.75	127.25	1	5	2	7	21.37	12.79	26.39	40.07
7647	100.25	3.6	18.85	2.4	7.12	32.15	2.24	17.61	10.63	77.75	124.5	1	9	1	7	23.63	12.41	37.81	68.22
DZLR24	98.2	5.6	19.9	2.3	8.48	37.7	3.47	49.65	9.375	77	123.75	1	7	2	7	24.81	13.64	32.53	52.59
7063	102.3	6.3	19.15	2.3	9.5	32.5	4	29.22	14.38	70	122	2	3	2	7	27.26	14.58	35.07	65.71
7581	97.1	5.35	19.1	2.6	8.88	33.9	3.21	44.63	9.375	79.75	127	1	5	2	7	22.73	12.65	31.57	73.64
242784	101.65	6.35	19.1	2.3	8.71	34.95	2.35	28.28	10.63	71.75	120.5	2	5	2	7	28.85	13.94	45.39	52.05
DZLR34	102.9	6.4	19	2.3	9.17	33.2	4.12	26.21	15	75.25	122.75	1	7	2	7	22.8	13.65	30.62	61.76
DZLR8	91.025	5.15	18.2	2.9	7.41	34.93	3.18	33.25	10	73	122.5	1	9	1	7	25.2	12.42	42.09	78.5
5515	96.55	5.65	19	2	7.95	34	2.21	26.2	8.75	77.25	119.25	1	7	2	3	24.21	12.67	33.4	54.2
DZLR7	104.3	5.45	18.9	2.4	9.08	31.2	3.55	30.76	13.13	74.5	124.5	1	5	2	7	31.62	14.35	49.65	82.41
242781	109.95	4.4	19.58	2.5	7.6	43.7	3.3	26.06	14.38	74.25	124.75	3	7	2	7	26.26	13.46	36.65	54.38
7641	96.25	8.15	18.8	2.4	9.33	34.25	4.58	54.51	9.375	74	118.75	3	7	1	7	26.36	12.53	40.77	84.97
5600	93.75	5.95	19.5	2.4	8.76	35.45	3.19	33.61	11.25	76.25	124	2	7	2	7	28.92	13.77	46.29	73.54
DZLR14	91.6	6.25	19.25	2.6	8.83	27.6	2.81	32.66	9.375	80.25	127.5	1	7	3	7	22.79	12.06	32.59	62.06
5434	104	4.9	19.3	2.1	8.6	37.15	4.14	60.82	6.875	74	124.25	1	9	2	7	23.21	12.84	28.63	61.61
6987	101.05	5.95	18.25	2.5	9	39.9	3	28.26	11.25	74.5	123.5	1	9	1	3	23.36	12.07	47.2	58.92
7007	99.95	6.45	18.6	2.1	8.34	34.93	4.51	39.83	11.25	73.25	125.5	2	7	2	7	22.5	12.55	24.19	51.68
DZLR22	99.725	5.3	18.3	2.3	8.58	33.48	2.04	27.1	8.75	76.25	126.75	1	3	3	3	27	13.27	44.56	77.62
243703	93.75	5.25	19.1	2.2	9.27	28.45	2.63	29	9.375	80	125.25	1	5	3	7	26.57	13.94	40.62	52.01
5760	111.6	5.7	19.05	2	7.74	35.45	3.48	46.04	8.75	73	122.5	1	7	2	7	25.41	12.26	36.1	59.85
DZLR31	96.2	6.45	18.5	2.3	9.5	30.05	3.29	24.17	13.13	75.5	126	2	7	3	7	22.06	12.01	39.24	73.58
6975	113.05	5.95	18.35	2.2	8.61	36.95	3.57	27.62	12.5	70.75	124	2	7	1	7	25.79	12.59	44.41	82.87
7050	97.95	6.35	18.6	2.4	7.89	35.95	4.32	40.3	11.25	73	121.5	2	5	2	7	27.53	13.22	37.96	71.8
5635	107.05	7.3	18.35	2.8	9.18	36.1	3.86	34.47	11.25	76	123.5	1	3	2	7	28.73	13.51	43.58	73.01
7532	101.65	6.7	19	2.3	8.95	34.8	3.23	25.62	13.75	77.5	125.25	1	5	2	7	23.16	12.12	31.16	64.89
5182	98.9	7.95	18.05	2	10	36.2	3.32	29.73	15	76.5	125.75	3	5	2	7	29.31	14.66	42.19	72.79

5057	102.8	6.05	18.1	1.875	8.83	33.75	4.04	49.83	10	71.25	124.5	3	5	2	7	29.26	14.55	39.2	52.37
7880	109.9	6.05	18.6	2.2	9.24	36.05	2.35	32.51	7.5	74.75	123.75	1	5	2	7	28.94	14.95	39.31	63.05
DZLR17	96.75	4.7	18.8	2.4	8.9	37.35	3.4	28.21	11.88	74.75	121.5	2	5	3	7	31.99	14.49	54.64	91.66
7626	106.05	4.9	18.15	2.2	8.8	30.95	3.26	26.2	15	69.75	118	1	7	2	7	29.99	14.64	43.55	61.08
7060	111.75	4.8	18.3	2.2	9.01	37.3	4.99	33.54	17.5	71.25	120	2	5	2	7	31.06	13.67	49.65	92.03
226694	105.6	5.8	18.35	2.2	8.88	36.2	3.77	28.82	12.5	75.25	125.75	1	7	2	7	30.07	13.91	48.31	80.16
242789	91.8	4.85	18.5	3	7.07	33.88	4.72	28.29	15.63	72.75	120.75	1	9	2	7	23.92	13.06	29.17	34.62
DZLR9	102.75	6.5	18.25	3.1	7.64	28.75	2.57	23.22	11.25	72.75	122.75	1	9	1	7	24.49	12.73	43.29	72.04
DZLR13	88.95	3.85	17.65	2.6	7.7	27.95	2.36	27.43	6.875	70.25	122.75	1	9	1	7	24.52	11.87	37.28	61.93
6102	108.25	6.9	18.6	1.55	10.1	31.5	2.74	25.76	9.375	70.75	120.75	1	7	3	7	35.22	13.94	47.4	76.64
5181	102.2	6.45	18.35	2.3	8.63	34.55	3.23	27.97	10.63	73.25	125.75	1	5	2	7	25.5	13.75	36.98	62.35
5707	105.9	5.1	18.7	2.4	9.28	36.25	3.76	36.4	10	72.25	121	1	3	2	7	28.4	14.43	41.58	78.18
6971	96	5.9	18.15	2.6	8.82	33.35	2.16	19.72	11.25	74.75	123.75	1	5	3	3	31.41	13.95	51.61	90.05
5169	103.85	6.2	18.6	2.2	7.62	34.8	3.96	36.32	12.5	73	123.25	1	5	2	7	26.26	13.56	41.29	61.21
6914	108.5	6.9	18.2	2.4	8.87	41.55	4.6	35.44	14.38	70.5	121	1	8	3	7	28.31	12.82	43.72	81.81
5149	99.85	8.5	18.5	2.1	9.13	35.15	3.95	29.04	15	72.5	122.25	1	3	2	7	29.64	13.96	50.46	88.01
222427	97.2	5.8	18.2	2.8	8.74	34.65	2.07	33.96	11.88	73.25	122.25	1	7	3	3	26.34	13.41	42.76	75.83
DZLR11	103.05	5.3	17.45	2.1	6.31	38.95	1.33	28.56	5.625	73.25	125.25	1	9	1	7	21.61	11.75	40.54	60.2
5729	103.25	7.25	17.63	2.4	9.5	37.55	2.39	29.32	8.75	75.5	125	2	7	2	7	26.58	12.96	31.47	50.19
DZLR6	100.3	6.95	18.7	2.2	9.81	38.25	3.96	36.38	11.88	75.5	125	1	5	2	7	30.06	14.86	44.08	79.02
5534	103.6	6.25	18.65	2.9	8.38	32.35	3.26	34.95	12.5	74.25	121.25	2	7	3	7	24.38	11.98	38.29	67.82
5638	96.2	4.3	18.95	2.3	7.65	35.05	4.75	21.92	22.5	79.25	123.5	1	7	1	7	26.27	12.74	45.65	91.03
5510	103.3	4.5	18.39	2.1	10.1	39.35	4.28	39.9	11.25	68.25	121	2	5	3	7	28.86	13.47	47.29	85.61
239711	104.725	7.6	18.85	2.6	9.65	36.2	5.97	71.77	8.125	67	119.25	2	7	2	7	28.19	13.53	37.68	42.69
DZLR40	91.85	7.3	17.5	2.5	8.98	30.6	1.94	21.75	8.125	79.5	128.5	1	1	2	7	26.75	13.57	28.05	50.48
DZLR2	103.65	5.3	18.3	2.1	8.05	29.55	2.53	25.4	9.375	71.25	120.25	1	9	1	3	27.21	13.46	46.42	87.65
DZLR26	103.35	5.85	18.25	2.1	9.64	36.95	3.55	31.71	12.5	78	127.5	2	5	2	7	26.44	14.23	39.39	52.78
5581	108.3	6.2	18.65	2.2	10.4	30.4	3.75	32.89	11.88	76.25	122.75	1	5	2	7	25.51	13.97	33.61	66.98
7664	101.8	5.7	18.8	2.6	7.26	30.95	3.6	26.33	14.38	79.75	125.25	2	7	3	7	25.96	13.05	42.14	79.51
7629	103.55	5.15	18.5	2	7.45	33	3.85	48.26	9.375	73.75	126	2	5	2	7	27.84	13.85	36.63	35.1
7825	97.95	7.4	18	2.9	8.55	30.45	3.37	32.13	14.38	72.5	121.75	1	7	3	7	27	13.18	39.07	77.06
6982	101.2	4.9	17.9	2.5	8.28	30.85	4.45	42.78	10.63	71	122	2	5	2	7	31.05	14.17	44.71	71.66

7015	102.85	4.7	17.5	2.675	9.09	35.65	4.62	38.81	12.5	67.75	117.5	2	7	2	7	27.17	13.86	37.35	58.1
7578	99.525	6.85	18.1	2.4	9.13	33.35	2.37	36.6	10	75.75	122.5	1	7	3	3	26.98	12.93	40.65	78.44
242790	96.25	6.05	17.8	2.775	9.03	34.45	2.87	24.7	11.88	72.5	124.25	2	7	3	7	28.95	13.59	43.66	57.76
243717	107.25	5.9	17.2	2	8.34	40.4	2.39	17.71	14.38	77	128	1	9	2	7	25.93	13.82	33.73	36.74
5373	95.95	6.25	18.2	2.7	8.34	32.35	2.3	26.4	9.375	79	127.25	1	5	3	7	23.32	12.15	33.38	66.38
5470	107.7	6.15	17.5	2.9	7.68	28.2	2.95	40.94	8.75	72.75	119.5	3	7	2	7	25.07	12.86	39.8	66.58
5576	88.25	4.95	17.4	2.8	7.97	32	2.44	26.23	10	75	126.75	1	7	2	7	26.41	13.57	34.76	54
7507	105.7	5.3	17.6	1.9	7.78	34.15	2.8	35.71	9.375	74.5	124	2	5	1	7	28.26	12.58	42.51	54.3
DZLR4	98.15	6.45	18.35	2.4	6.74	30.65	2.46	25.22	10	75.25	124.5	1	9	1	7	25.11	13.5	47.63	78.55
7009	106.65	6.25	18.15	2.3	8.65	34.03	3.77	59.49	7.5	69	120	2	5	3	7	28.04	13.15	39.38	80.97
242783	97.15	7.25	17.36	2.4	7.98	39.35	2.3	18.45	12.5	77.5	124	1	9	2	7	25.39	13.62	37.55	45.98
Werer	85.65	3.325	18.15	3.8	6.4	39.35	3.3	24.96	13.13	65.5	122.75	1	9	1	7	28.27	13.34	56.53	91.75
DZLR32	106.5	4.75	17.9	2.5	8.93	32.6	2.82	20.95	15	76.5	125	2	5	3	7	23.24	12.25	41.39	80.37
7242	102.85	4.85	17.85	2.2	9	34.15	4.13	30.77	15.63	71	122.25	2	5	3	7	28.55	13.46	41.8	78.13
7673	111.55	6.2	17.8	2.2	8.96	35.4	3.11	43.14	7.5	69.5	119.75	1	5	2	7	25.5	12.63	35.62	52.36
Arendato	92.9	5.75	18.05	2.5	6.9	27.28	2.03	24.96	7.5	75.25	122.25	1	9	1	7	26.81	13.36	49.55	85.59
DZLR39	91.65	5.85	17.9	2.3	7.27	32.2	3.34	37.62	8.75	74.5	121.75	1	9	1	3	27.27	13.28	48.45	86.31
5618	101.25	5.95	18.1	2.3	6.53	30.75	3.28	17.94	19.38	76	126	1	9	3	7	28.1	13.96	50.77	86.45
5584	98.45	5.95	17.15	2.35	8.09	36.35	3.38	35.6	10.63	77.25	130	3	3	2	3	26.69	13.44	39.24	65.47
5369	98.2	5.7	18.31	2.5	7.71	30.55	2.48	26.51	10	81.25	122.75	1	9	3	7	23.12	12.25	35.65	62.93
5492	89.25	5.85	18.33	2	5.57	32.85	2.5	33.22	9.375	79.5	127.5	1	9	2	3	27.67	13.63	51.58	82.19
7571	91.1	5.05	18.25	2.5	8.48	34.5	2.9	63.76	5.625	70.5	124.75	1	7	2	7	22.2	12.44	25.77	29.94
243706	111.65	5.35	17.4	2.2	7.43	29.8	2.78	15.61	16.88	74.5	122.25	1	5	3	7	27.62	13.53	44.86	70.54
5441	101.51	6.3	17.8	2.3	8.88	33.85	3.78	24.97	15	73.75	123.75	2	3	3	7	23.97	12.47	44.28	78.09
5913	103.9	4.35	17.9	2.3	6.41	28.48	2.79	19.36	15	69	118.5	1	9	2	7	26.49	12.89	50.08	64.66
DZLR5	104.65	6.15	17.95	1.9	8.78	35.45	3.03	37.61	11.25	73.5	122.25	3	7	2	7	27.77	13.08	38.89	72.4
7827	102.9	5.55	16.9	3	8.38	31.5	3.49	34.67	10.63	75.5	124.5	1	7	3	7	23.32	11.94	33.64	72.86
5172	97.2	5.55	17.05	2.2	8.62	32.3	3.06	38.91	9.375	74.75	122.5	2	5	2	7	26.18	13.53	36.45	56.16
7195	101.2	6.2	17.8	2.2	8.79	35.7	2.79	21.05	13.13	78	126.75	1	7	2	7	25.3	13.18	32.6	64.68
238891	103.95	7.25	18	2	8.5	37	3.97	36.35	11.25	73.75	124.75	1	5	3	7	35.26	14.56	58.42	75.4
Tob66	101.95	5.65	18.15	3.9	7.65	41.5	4.7	36.3	13.13	73.75	128.5	1	9	3	7	28.35	13.15	47	82.6

DZLR20	93.4	6.1	17.35	2.9	9.23	34.65	2.84	21.64	13.75	76.75	129	2	5	3	7	28.42	13.79	49.97	72.68
5342	99.45	5.85	17.35	1.8	8.69	34.6	1.49	13.55	11.88	79	126.5	1	5	2	7	30.98	14.53	45.41	73.88
7078	106.85	8.1	17.5	2.2	9.45	37.5	4.52	34.4	13.75	73.5	123	1	5	2	7	27.02	14.09	37.23	72.24
7209	105.325	5.15	17.35	1.8	7.42	37.9	3.88	25.87	15.63	79.25	127.25	1	9	2	3	24.48	12.26	43.74	88.51
5183	92.5	6.75	17.65	2.1	8.19	34.1	3.86	47.59	10.63	70.75	121	2	7	2	7	24.91	12.39	30.61	72.3
DZLR35	93.6	5.9	17.6	1.7	7.73	34.9	3.97	46.5	11.88	69.5	123.25	1	5	3	7	26.29	12.82	46.03	74.87
6968	102.3	6.4	17.51	2.65	7.63	35.05	2.46	26.91	8.75	75.25	125.25	1	7	3	7	25.21	12.53	32.46	67.61
7313	107.45	5.3	18.05	2.5	9.4	37.2	3.11	43.16	7.5	68	120.5	2	5	3	7	34.61	14.78	52.36	73.57
7798	102.25	6.15	17.25	2.5	8.61	32.55	3.12	49.56	8.125	76.25	125.5	2	5	1	7	22.36	12.44	36.54	66.66
242786	90.65	6.65	17.25	2.4	8.46	34	2.92	28.52	10	72.5	122.25	1	9	2	7	24.67	13.32	37.53	50.82
DZLR12	100.95	6.95	17.5	2.3	7.82	34.05	2.4	23.82	10	79.25	128.25	1	3	3	3	22.88	12.13	33.76	64.95
DZLR19	99.95	5.9	17.7	2.2	8.43	29.8	2.62	16.87	15.63	78.75	124	1	3	3	3	24.53	12.87	38.99	73.47
DZLR1	100.6	5.35	17.85	2.6	8.25	34.35	2.98	21	13.75	71.75	125.25	1	3	3	7	28.45	14.16	52.22	63.51
DZLR27	82.55	5.5	17.5	4.1	7.55	36	3.59	48.46	7.5	68.5	123	1	9	3	7	30.26	13.95	57.61	91.32
7014	106.8	6.45	17.25	2	10.1	35.2	3.86	36.22	10.63	72.5	123.5	2	5	3	7	26.58	13.69	34.21	47.43
5174	101.65	5	17.25	2.1	6.53	30.5	3.62	37.88	10.63	73.75	122.5	1	9	1	7	24.94	12.97	46.71	86.21
5504	109.4	5.35	17.4	2.2	8.63	32.75	3.16	42.07	13.75	79.75	126.75	2	5	2	7	31.65	15.29	42.65	66.3
5666	98.8	5.9	17.44	1.875	9.78	37.75	3.61	29.96	12.5	73.5	122.5	1	3	2	7	28.44	14.66	44.18	77.5
DZLR9	92.25	7	17.25	2.5	8.98	35.9	2.91	25.83	11.25	74.25	126.75	2	5	2	7	26.54	12.8	40.76	81.37
DZLR29	102.7	7.35	17.85	2.1	9.31	32.8	2.84	36.16	8.75	76	126.5	2	5	2	7	26.68	14.01	39	56.58
5609	103.1	5.65	17.75	2.3	9.57	31.6	2.83	35.03	7.5	78	125.5	1	7	3	7	28.54	13.12	43.59	64.59
7020	107.5	7.5	17.45	2.3	9.75	30.88	3.6	32.99	10.63	72.25	118.75	2	7	2	7	32.32	14.33	51.05	94.56
7378	101.9	6.5	17.05	2.1	7.6	38.2	3.62	24.04	16.25	73	122.25	3	7	3	7	23.77	11.89	38.43	68.37
7075	101.6	4.95	17.4	2.1	9.01	33.15	4.11	30.92	13.75	67	117.75	1	7	2	7	28.53	14.39	39.07	42.44
DZLR37	92.05	5.7	16.7	2.3	8	35.05	2.72	25.24	10.63	68.75	121.5	2	7	2	7	19.94	11.48	30.83	36.02
DZLR23	94	5.3	17.25	2.4	7.94	35.25	2.86	35.54	9.375	73.5	120	3	7	3	7	27.4	12.93	35.91	70.32
5197	103.65	5.7	16.75	2.3	9.01	35.6	3.69	52.31	8.75	69.75	117.5	2	5	3	7	22.3	11.57	39.56	82
5214	94.95	6.85	16.65	2.4	8.8	30.05	2.75	40.01	10	76.75	126.25	1	5	2	7	27.95	13.46	44.03	76.88
6955	92.25	7.9	17.85	2.65	7	28.78	2.82	35.49	8.125	74	123.75	1	7	2	7	25.16	15.9	29.47	28.48
6933	102.8	6.55	16.8	2.225	8.46	34.4	2.61	28.63	8.75	72.25	121.25	1	5	2	7	27.41	13.84	41.98	58.22
5171	93.95	6.95	17.45	2	8.05	37.75	2.87	35.54	9.375	77.5	126.25	2	5	2	7	24.43	12.25	32.68	69.9

Yerer	81.85	6	17.9	3.7	8.15	40.3	4.54	52.81	9.375	69.25	124.25	1	9	3	7	29.07	13.69	55.64	88.85
5491	99.7	6.05	16.75	2.3	8.05	35.8	3.9	35.6	11.25	77.75	126.75	2	5	2	7	24.39	14.77	29.17	44.49
7150	106.85	5.5	17.85	2.2	6.21	25.1	2.39	16.65	13.13	75.5	125.75	1	9	1	7	27.21	14.14	51.09	81.94
7477	95.75	5.85	17.3	2.1	9.2	29.65	2.25	19.3	12.5	76.25	128.75	1	5	2	7	20.99	12.3	28.45	30.76
DZLR36	109.95	5.55	18.05	3.3	8.51	44.9	4.38	72.21	6.25	64.75	122.5	1	9	2	7	23.74	11.84	47.49	73.18
5216	99.7	6.25	18	3.2	8.39	29.15	3.71	24.27	13.75	73	124	1	5	2	7	28.58	13.8	45.38	76.62
5591	99.35	7	16.5	2.2	7.88	34.8	2.76	25.63	15	72.5	125	3	5	2	7	22.71	13.2	39.69	50.77
235051	105.5	5.3	16.8	2.1	7.93	36.13	2.77	24.63	10.63	71.5	123.5	2	7	1	7	31.01	14.67	49.72	79.7
DZLR25	106.05	8.4	17.7	2.2	6.88	36.2	2.5	22.55	12.5	69.75	121.25	3	7	2	7	25.42	12.37	42.34	82.66
5142	102.5	5.9	17.2	2.5	8.01	30.35	2.76	30.57	9.375	77.25	124.75	2	9	2	7	19.94	11.24	19.48	46.44
7069	100.45	4.85	17.25	2.3	8.46	35.05	3.89	38.65	10	67.5	121.5	2	7	2	7	28.73	14.3	37.6	53.62
7713	99.15	5.1	17.5	2.2	6.95	31.1	3.21	31.16	10.63	77.25	122.5	1	3	2	7	24.91	13.36	36.61	37.51
7823	97.4	6.05	16.75	2.5	7.09	36.2	3.1	37.12	10	77.75	125.5	2	0	3	7	24.42	12.31	38.78	75.63
Gerardo	91.25	6.45	17.15	3.5	9.02	42.35	4.18	51.38	13.75	70.25	119.5	2	9	3	7	24.47	12.67	46.13	75.87
5728	98.45	7.75	17.1	2.375	9.62	36.5	4.52	32.48	13.75	72.25	118.5	2	3	2	7	27.94	15.44	42.65	74.68
7076	96.75	5.65	17.2	2.1	9.46	36.25	4.12	64.99	7.5	72.5	122.5	1	5	2	7	28.23	14.35	39.92	54.61
7832	103.7	6.25	17.8	2.3	7.27	32.25	3.25	28	13.13	76.25	125.75	1	5	2	7	24.76	12.72	52.48	76.22
DZLR15	107.4	5.95	16.75	2.7	9	34.45	3.64	32.19	12.5	77.5	125.5	1	3	2	7	27.31	13.08	43.3	87.71
5044	104.85	6.6	17.25	2.4	9.88	34.55	2.35	27.12	11.25	76.25	126.75	1	0	2	7	32.37	14.12	51.46	62.7
7004	104.05	5.2	17	2.5	9.04	31.45	3.8	36.65	11.25	69.5	120.75	2	7	2	7	25.02	14.23	31.98	45.2
5140	110.3	7	16.7	1.9	9.88	35.3	3.83	28.53	13.75	74.5	121.25	1	5	2	7	26.4	13.72	44.02	77.8
7568	101.175	5.25	17.1	2.7	8.13	27.05	2.82	22.62	11.88	73.75	123	3	7	3	7	25.62	13.4	31.1	58.85
243700	100.3	8.2	17.5	2.5	7.8	32.13	3.26	33.21	10	69	120.75	1	7	2	7	29.4	14.43	37.31	54.17
7464	98.65	5.65	17.5	2.2	8.24	35.13	2.7	29.2	11.25	74	129.5	1	5	2	7	24	12.75	25.33	45.04
5893	99.5	6.05	16.75	2.4	7.89	39.05	3.77	40.35	10	78	123.75	1	5	1	7	23.92	12.79	47.48	80.99
7056	100.4	7.2	17.65	1.9	9.19	32.8	4.56	32.57	16.25	73	122.5	2	5	2	7	30.39	13.84	44.11	83.29
Alemtena	89.1	4.9	17.55	4.5	7.43	31.1	3.23	26.82	11.88	67.5	118	1	9	1	7	23.98	12.58	46.18	80.65
7003	103.05	6.15	17.9	2	7.93	34.7	3.66	19.02	18.13	69.75	123.25	1	5	3	7	29.9	14.14	46.08	65.64
7999	95.1	6.7	16.6	2	7.48	30.9	2.19	43.03	8.125	76.75	124.5	2	7	2	3	29.33	14.83	43.16	50.32
Tesfaye	83.75	5.85	17.35	3.7	5.94	33.4	3.92	62.8	8.125	64	118.25	1	9	3	7	27.87	13.68	57.35	86.72
DZLR28	93	6.15	17.2	2.2	8.01	25.75	2.44	17.3	14.38	73.25	124.5	1	7	3	7	27.39	14.49	42.22	52.5

7826	93.25	6	16.8	2.5	7.4	26.6	2.42	30.68	7.5	75	124.25	1	5	2	7	21.98	12.18	30.84	30.5
5052	117.86	5.8	17.65	2.7	8.39	28.48	3.62	32.67	13.13	68.5	120.25	1	5	2	7	26.58	13.44	42.31	77.26
5168	92.15	5.4	16.7	2	8.6	32.9	3.1	22.47	13.75	73	122.5	2	7	3	7	29.21	13.71	44.41	77.66
7295	101.9	6.5	17.15	1.9	8.43	38.15	3.69	28.48	15	70	121	1	5	2	7	26.28	12.74	40.89	67.86
DZLR33	93.85	5.85	16.55	2.2	8.65	31.75	2.67	54.61	6.25	78.25	125.25	1	5	2	7	21.57	12.04	32.58	33.66
242793	112.65	5.75	16.95	2.2	6.38	26.85	2.23	21.79	11.88	68.25	123.75	1	9	3	7	26.54	13.4	46.72	53.34
7762	101.05	5.9	16.25	1.775	8.71	41.1	2.8	30.98	12.5	80.5	127.5	1	7	3	7	24.31	12.13	40.58	75.64
Quamy	99.75	4.55	17.6	3.7	7.12	41.9	4.27	41.71	13.75	71.25	125.5	2	9	3	7	27.61	13.21	45.61	85.98
7690	91.95	6.45	16.85	2.4	6.47	34.2	2.61	23.71	10.63	76	125.25	1	3	1	7	24.24	12.71	40.81	67.53
Ld-357	100.85	5.95	16.65	2.8	6.89	38.65	2.93	30.94	10.63	71.25	125	2	9	1	7	25.55	14.23	50.77	67.07
Ejersa	85.1	6	17.15	3.75	6.87	37.85	2.3	26.97	10.63	65	122	1	9	3	7	29.31	13.63	58.87	93.04
DZLR18	96.45	5.2	16.6	2.7	8.15	27.05	1.83	50.64	3.75	68	117.5	2	5	2	7	22.85	13.1	28.84	40.23
Foka	107.075	5	17.18	3.3	7.05	41.75	4.67	46.74	10	72.75	123.5	1	9	3	7	27.18	12.96	49.47	89.96
242792	95.9	5.85	15.85	2.1	7.66	36.3	3.98	61.53	7.5	74.75	123.25	2	3	2	7	17.89	11.14	21.36	54.03
242782	106.55	7.7	16.3	2	7.95	36.75	2.7	32.06	8.75	70.5	124.25	2	5	2	7	24.47	13.01	37.34	57.1
243701	94.7	6.2	16.2	2.6	8.97	29.25	2.51	24.65	10	78.5	128	1	5	2	7	25.25	12.77	26.02	47.54
7666	101.25	6.1	16.05	2.4	7.18	38.25	3.84	29.51	12.5	66.5	121	2	7	3	7	24.04	11.91	43.18	81.69
Malefia	91.6	5.85	16.55	4	6.93	26.45	1.73	34.05	6.25	73.25	125	1	9	1	7	26.86	12.79	52.23	57.83
Mangudo	84.85	5.2	17.05	3.5	5.88	41.1	4.15	48.57	10	66.25	119.5	1	9	3	7	27.82	12.92	53.95	83.17
DZLR30	95.05	5.6	16.65	2.2	8.29	29.85	2.36	26.53	6.875	75.25	123.5	1	7	2	7	26.63	12.52	37.33	70.76
7683	100.6	5.7	16.9	2.5	8.55	30.55	2.77	24.2	11.88	75.75	125.25	1	9	2	7	30.33	14.15	37.12	62.41
DZLR16	96.725	7.25	16.95	2.4	8.73	28.75	3.64	27.43	15	63.5	117.5	1	3	3	7	26.44	13.37	45.57	69.44
DZLR10	107.55	7	16.79	3.6	6.98	44.65	4.45	30.76	14.38	61	117.25	1	9	1	7	23.17	11.87	47.67	87.51
5184	99.65	4.85	16.53	2.1	7.49	29.78	1.61	17.64	11.25	75	126.25	1	7	1	7	32.84	14.08	63.19	91.27
Fetan	89.3	5.1	16.8	4.3	6.93	29.85	3.73	30.48	12.5	69.25	122.75	1	9	3	7	24.67	12.83	48.05	87.79
DZ2018	81.45	6.8	16.25	4	7.33	35.95	2.55	62.83	4.375	69	122.5	1	9	3	7	28.28	13.35	54.41	79.44
5586	107.2	6	17	2.735	6.84	36.15	2.86	19.25	15.63	79.25	129.5	1	7	3	7	25.48	12.87	45.9	90.84
7191	102.25	6.4	16.34	2.3	7.18	35.3	3.69	43	9.375	74.75	124.75	1	7	3	7	25.12	12.78	41.46	70.42
Mukiye	86.2	5.7	16.7	3.7	6.63	40.75	2.67	37.01	7.5	64.5	119.25	1	9	3	7	25.06	12.56	48.79	81.75
5554	91.35	6.35	16.95	2.6	6.63	31.7	2.78	23.75	12.5	78	125	1	9	3	3	25.62	12.99	46.56	81.05
Selam	102.75	5.7	16.55	3.8	6.63	43.3	3.61	24.88	14.38	71.5	124.5	1	9	3	7	26.85	12.98	54.63	89.1
Asasa	106.15	7.95	16.9	3.5	6.19	39.4	4.01	40.97	11.25	62.25	119	1	9	3	7	25.45	12.93	53.87	93.35

Bekelcha	88.7	5.95	16.9	3.5	7.05	36.25	3.69	41.69	9.375	68.25	118	1	9	3	7	26.37	12.79	47.53	80.69
DZLR38	96.7	6.3	16.05	2.35	7.63	26.75	2.81	23.92	11.25	68.75	118.25	2	5	2	7	26.51	13.86	32.29	40.85
Ginchi	103.2	5.95	16.15	3.7	6.88	41	4.04	54.04	7.5	62	120.25	1	9	3	7	24.89	11.97	50.6	86.21
5166	96.65	5.7	15.6	3	7.5	23.6	2.68	24.76	10	74.5	124.75	1	9	3	7	21.16	12.58	36.29	68.63
Boahi	111.5	6.25	15.95	3.325	6.2	39.78	3.79	38.15	10	74.75	121	1	9	3	7	27.17	13.2	53.12	88.29
214370	101.35	6.6	15.75	1.6	8.94	37.3	2.21	17.56	13.13	78.25	124.5	1	9	1	7	31.48	14.83	62.89	98.18
Cocorit/71	82.75	4.55	15.2	3.8	6.97	38.28	3.82	34.33	13.13	65.75	119.75	1	9	1	7	24.91	15.2	46.86	72.03
5465	106.5	5.675	17.93	2.85	7.97	37.81	4.45	39.69	12.5	72.25	128.5	1	9	2	7	31.91	15.05	45.94	57.97
Mean	99.86	5.99	18	2	8.2	34.24	3.25	33.2	11.23	73	123	1	5	2	7	26.49	13.3	41.57	68.4

NB: GY= grain yield (t/ha), BY= Biomass yield (t/ha), TKW= thousand kernel weight, PH= plant height (cm), TPP= number of effective tillers per plant, SpSp= number of spikelets per spike, SDpSp= number seeds per spikelet, SPL- spike length (cm), HI= harvest index, DB= days to booting, DM= days to maturity, GC= glume color, SD=spike density, SC= seed color, Aw=awndness, GC(%) = gluten content (%), MC = moisture content (%), PC- protein content (%), ZI= Zeleny index (ml), VITP= vitreous percentage, DZLR= landrace from Debrezeit, DZLR1= ETDW/15DZ#, DZLR2= ETDW15DZ012,ETDW/15DZ070, DZLR3= ETDW15DZ073,ETDW/15DZ009,……., DZLR4 = ETDW/15DZ055,MCD3-1, DZLR5=ETDW/15DZ#12,ETDW/15DZ#21……,DZLR6=ETDW/15DZ#50,ETDW/15DZ#52,……, DZLR7 = ETDW/15DZ#39, DZLR8 = ETDW/15DZ#43ETDW/15DZ#068,….., DZLR9 = ETDW/15DZ043,DZDW17010.8, DZLR10 = ETDW/15DZ#044, DZLR11 = ETDW/15DZ014, DZLR12 = ETDW/15DZ#4,ETDW/15DZ#20……..., DZLR13 = ETDW/15DZ#21ETDW/15DZ#31, DZLR14 = ETDW/15DZ036,ETDW/15DZ063, DZLR15= MCD12-30, DZLR16 = MCD4-32, DZLR17 = DZDW17007, DZLR18 = MCD14-3,MCD11-28,MCD11-36,……., DZLR19 = DZDW176,7,10,DZDW17005,DZDW17006,07, DZLR20 = MCD8-44, DZLR21 = MCD5-4,MCD5-25,MCD6-44,……, DZLR22, MCD3-19,ETDW/15DZ#34,ETDW/15DZ#35, DZLR23 = MCD12-2,MCD12-38,MCD14-26,……., DZLR24 = MCD13-8,MCD14-29, DZLR25 = MCD4-4, DZLR26= MCD10-11, DZLR27 = DZDW17002, DZLR28 = MCD11-7,MCD11-25,MCD13-5,MCD14-19, DZLR29 MCD10-3,MCD10-16,MCD10-44…… DZLR30 = MCD5-18, DZLR31 = MCD3-27, DZLR32 = MCD3-14, DZLR33 = MCD7-35, DZLR34 = MCD14-29, DZLR35 =MCD11-20, MCD8-2, ETDW/15DZ019,……, DZLR36 = MCD3-15, DZLR37=MCD5-35,ETDW/15DZ004, DZLR38 = MCD4-32-2, DZLR39 = MCD4-13,MCD8-6,MCD11-18,……., and DZLR40 =DZDW17004,DZDW1702,04.

Annex 4 Summary of the putative QTLs identified across tetraploid wheat chromosomes for agronomic and grain quality traits combined across locations and years

Putative QTL	chr	Map position (bp)	Position* (Mbp)	Flanking markers	Position** (Mbp)	Effect	SNP	alleles	R2	Reports
qGY1	3B	13884562	9.834619	982968|F|0-12:C>T-12:C>T/1101924|F|0-39:C>G-39:C>G	17.968757	0.528617	19:T>A	T/A	19.90621	Sukumaran *et al*., 2014; Lopes *et al*., 2014; Neumann *et al.*, 2011; Jamil et al., 2019; Maccaferri *et al.*, 2008; Roncallo *et al*., 2017
qGY2	6B	590911770	586.66169	1126049|F|0-26:A>C-26:A>C/978845|F|0-64:A>G-64:A>G	597.557368	-0.80179	18:A>T	A/T	43.7197	Ain *et al*., 2015; Rahimi *et al*., 2019 and Neumann *et al*., 2011
qAwn1*	4B	504183168	495.58499	1090973|F|0-54:C>T-54:C>T/3961293|F|0-19:C>T-19:C>T	512.797498	-1.61935	6:A>G	A/G	46.53487	PleotropicNovel
qAwn2*	4B	504183168	495.58499	1090973|F|0-54:C>T-54:C>T/3961293|F|0-19:C>T-19:C>T	512.797498	0.694008	6:A>G	A/G	46.53487	PleotropicNovel
qDB1	1A	379455983	375.06514	5581922|F|0-13:G>C-13:G>C/1054139|F|0-63:A>G-63:A>G	384.370732	-2.56893	67:C>A	C/A	8.765336	Mehrabi *et al*., (2020)
qDB2	2B	202155886	198.15072	3026637|F|0-26:A>T-26:A>T/985339|F|0-27:A>G-27:A>G	206.689407	1.235177	64:G>A	G/A	2.286672	Mengistu *et al*. (2016)
qDB3	5A	18572611	12.008969	1215511|F|0-44:T>C-44:T>C/1044409|F|0-32:G>C-32:G>C	22.72033	1.083248	8:T>C	T/C	2.082928	Mehrabi *et al*., (2020)
qDB	5B	703291898	699.39461	1228874|F|0-14:G>A-14:G>A/4005191|F|0-34:A>G-34:A>G	707.379766	-1.435805	30:C>T	C/T	2.540662	Mehrabi *et al*., (2020)
qDB4	6B	28685444	24.465523	3961067|F|0-32:T>C-32:T>C/983457|F|0-61:G>A-61:G>A	34.223583	2.356596	57:G>T	G/T	4.936379	Mehrabi *et al*., (2020)
qDB5	6B	142457155	134.8247	1026550|F|0-42:T>C-42:T>C/1074208|F|0-7:A>G-7:A>G	147.072924	-3.10883	18:C>A	C/A	12.32457	Mehrabi *et al*., (2020)
qDB6	6B	590911770	586.66169	1126049|F|0-26:A>C-26:A>C/978845|F|0-64:A>G-64:A>G	597.557368	2.871329	18:A>T	A/T	21.22428	Mehrabi *et al*., (2020)
qDM1	4B	535089105	529.56393	1089380|F|0-37:G>A-37:G>A/1089013|F|0-68:A>C-68:A>C	540.718898	1.951566	15:C>T	C/T	22.42127	Sukumaran *et al*., 2014
qDM2	6B	80781189	76.216654	2331384|F|0-53:A>C-53:A>C/4542807|F|0-19:T>G-19:T>G	84.992114	-1.49264	24:A>G	A/G	18.07309	Mwadzingeni *et al*., 2017
qGCW1	2B	153613201	148.86716	5323433|F|0-53:A>C-53:A>C/1021401|F|0-10:C>G-10:C>G	157.739168	0.327558	27:G>A	G/A	18.10301	Novel
qGCW2	5B	356684358	350.49757	12776212|F|0-20:T>C-20:T>C/987795|F|0-10:A>G-10:A>G	362.127576	0.280917	32:C>A	C/A	16.47112	Novel
qHI1	1A	368867751	362.22025	2325748|F|0-30:G>A-30:G>A/1087379|F|0-64:G>A-64:G>A	374.165853	-8.44013	23:G>C	G/C	33.93406	Neumann *et al*. (2011); Assanga *et al*., 2017
qHI2	6B	473085307	465.67351	1863340|F|0-29:G>A-29:G>A/1054930|F|0-60:C>A-60:C>A	477.696003	4.312003	56:G>A	G/A	11.06407	Ain *et al*., 2015
qPH1	2A	75591116	69.314494	992816|F|0-7:A>G-7:A>G/1105695|F|0-15:C>G-15:C>G	80.537333	-5.826	46:G>C	G/C	12.37518	Ain *et al*., 2015
qPH2	2A	746611005	742.71464	4993789|F|0-5:C>A-5:C>A/1093498|F|0-43:G>T-43:G>T	750.550903	-7.34834	67:A>G	A/G	19.44663	Ain *et al*., 2015
qPH3	2B	632449752	623.86873	7340829|F|0-12:G>A-12:G>A/1107071|F|0-54:T>C-54:T>C	636.543387	-0.77144	15:A>G	A/G	6.715567	Assanga *et al*., 2017; Mwadzingeni *et al*., 2017; Maccaferri *et al*., 2008;

qPH4	4B	390403261	386.49148	1061937|F|0-48:A>C-48:A>C/2289037|F|0-62:T>C-62:T>C	397.032617	-4.86192	20:G>C	G/C	3.68705	Tzarfati *et al*., 2014; Mengistu *et al*., 2016; Ain *et al*., 2015
qPH5	5B	562792380	558.57566	3020370|F|0-14:T>C-14:T>C/1072508|F|0-32:G>A-32:G>A	567.017032	-3.25247	11:C>A	C/A	3.521646	Jamil *et al*., 2019; Mengistu *et al*., 2016; Wang *et al*., 2019 Ain *et al*., 2015; Mwadzingeni *et al*., 2017; Rahimi *et al*., 2019
qPH6	5B	703899766	700.04396	1013370|F|0-19:C>T-19:C>T/1139663|F|0-9:C>T-9:C>T	708.325171	3.305213	35:T>G	T/G	4.826288	Ain *et al*., 2015; Mwadzingeni *et al*., 2017; Rahimi *et* al., 2019
qPH7	7B	37247484	32.549004	1099958|F|0-57:G>A-57:G>A/1278515|F|0-6:T>C-6:T>C	41.428605	2.30551	36:C>T	C/T	1.701493	Kumar *et al*., 2020; Mwadzingeni *et al*., 2017; Tzarfati *et al*., 2014
qPSDC1	3A	673212466	667.79456	1210159|F|0-12:A>C-12:A>C/12698509|F|0-11:A>G-11:A>G	678.283587	-0.63864	18:T>A	T/A	6.023172	Alemu *et al*. (2020)
qPSDC2	7B	639474504	633.85822	2276419|F|0-50:T>C-50:T>C/4004583|F|0-22:T>C-22:T>C	643.664111	0.093092	47:T>C	T/C	16.75006	Alemu *et al*. (2020)
qRSDC	7B	656649140	652.37414	2279311|F|0-44:T>C-44:T>C/1099572|F|0-35:T>C-35:T>C	660.616677	0.358609	28:G>A	G/A	19.71539	Alemu *et al*. (2020)
qWSDC	3B	710224060	705.74302	3953020|F|0-23:G>T-23:G>T/12768830|F|0-66:C>A-66:C>A	715.858671	0.184519	14:A>C	A/C	36.13584	Alemu *et al*. (2020)
qSdPSp1*	2B	91084823	87.163601	2292727|F|0-44:T>C-44:T>C/1092239|F|0-47:G>T-47:G>T	95.769355	-0.32733	6:T>G	T/G	9.169342	Roncallo *et al*., 2017
qSdPSp2	5B	550320578	544.33911	2275681|F|0-40:A>G-40:A>G/100082073|F|0-64:C>G-64:C>G	554.523085	0.283962	6:T>C	T/C	21.76123	Rahimi *et al*. (2019
qSdPSp3	6B	705490559	701.16689	1153424|F|0-35:T>C-35:T>C/1291647|F|0-15:G>A-15:G>A	710.011435	-0.42883	17:C>G	C/G	48.33304	Novel
qSpDVD	2B	380821968	372.89502	100244487|F|0-60:T>C-60:T>C/1071174|F|0-68:A>G-68:A>G	392.617463	-3.38887	30:T>C	T/C	36.69704	Novel
qSpDVL1	1A	384370732	379.45598	978646|F|0-67:C>A-67:C>A/1125019|F|0-50:G>A-50:G>A	389.424444	0.201105	63:A>G	A/G	15.51887	Novel
qSpDVL2	1A	520852282	513.79784	4993889|F|0-6:T>C-6:T>C/3020449|F|0-43:A>G-43:A>G	526.692658	0.075065	16:G>A	G/A	2.257446	Novel
qSpDVL3	2B	762937294	759.17566	1190612|F|0-10:A>G-10:A>G/1108614|F|0-35:A>C-35:A>C	767.019751	-0.05589	41:T>A	T/A	6.02E-08	Novel
qSpDVL4	3A	657422205	652.93792	3385604|F|0-10:G>A-10:G>A/995571|F|0-20:C>A-20:C>A	661.927551	0.109921	26:G>C	G/C	5.18567	Liu *et al*., 2020;Halder *et al*., 2023
qSpDVL5	5A	657655271	657.650685	1218002|F|0-25:C>G-25:C>G/1129930|F|0-47:T>C-47:T>C	657.655337	0.14535	16:T>C	T/C	32.55087	Liu *et al*., 2020
qSpDVL6	6A	485391503	481.34116	1714673|F|0-64:C>T-64:C>T/1219825|F|0-14:C>G-14:C>G	489.229764	0.082964	14:C>T	C/T	0.187291	Liu *et al*., 2020
qSpDVL7	6A	532657424	525.94909	3935517|F|0-27:G>A-27:G>A/1107875|F|0-22:T>C-22:T>C	537.313492	-0.09578	61:C>T	C/T	30.02981	Ain *et al*., 2015
qTKW1	2A	42932966	388.6749	12770581|F|0-16:C>T-16:C>T/1159845|F|0-58:C>T-58:C>T	471.818337	1.778163	23:T>C	T/G	5.428667	Sara *et al*., 2020
qTKW2*	2B	91084823	87.163601	2292727|F|0-44:T>C-44:T>C/1092239|F|0-47:G>T-47:G>T	95.769355	-2.60604	6:T>G	T/G	30.48988	Assanga *et al*., 2017; Rahimi *et al*., 2019; Neumann *et al*., 2011; Ain *et al*., 2015; Jamil *et al*., 2019
qTKW3	3A	479775382	474.55477	2279238|F|0-47:C>T-47:C>T/1005033|F|0-11:G>A-11:G>A	483.743419	1.550066	38:A>G	A/G	3.90482	Ain *et al*., 2015
qVITP1	3B	532749194	527.27645	1057295|F|0-5:C>T-5:C>T/3064551|F|0-14:A>G-14:A>G	537.236924	10.77953	48:T>A	T/A	14.47718	Novel
qVITP2	5A	688984224	685.14162	2277988|F|0-20:T>C-20:T>C/2279526|F|0-19:G>C-19:G>C	692.832238	14.0834	17:G>A	G/A	13.62557	Sun *et al*., 2018
qVITP3	5B	491070306	486.28012	3029090|F|0-37:G>T-37:G>T/1218141|F|0-16:T>C-16:T>C	495.53758	9.760137	16:T>C	T/C	15.58107	Johnson *et al*., 2019; Sun *et al*., 2018
qZI	3A	673588291	669.55625	3939466|F|0-22:C>T-22:C>T/12698509|F|0-11:A>G-11:A>G	678.283587	-0.04279	18:T>A	T/A	37.34477	Novel

Remakk: * is to indicate pliotropic loci.

Annex 5 Functional annotation of genes identified within the QTL regions

Trait	genes	Protein coded	Chromosome
Grain yield	TRITD3Bv1G007350,TRITD3Bv1G007370,TRITD3Bv1G005230,TRITD3Bv1G005250, TRITD3Bv1G005980, TRITD3Bv1G006530, TRITD6Bv1G185720 and TRITD6Bv1G187460	F-box protein	3B, 6B
GY,HI	TRITD6Bv1G144950, TRITD6Bv1G145040	Expansin	6B
GY	TRITD3Bv1G007970, TRITD3Bv1G008060, and TRITD3Bv1G008000	Expansin-A24	3B
HI	TRITD1Av1G136830,TRITD1Av1G136660,TRITD1Av1G136580 ,TRITD1Av1G136420,and TRITD1Av1G136610	Expansin-B4,-B3,-B6,–B2	1A
DB	TRITD6Bv1G009410, TRITD6Bv1G009330, and TRITD6Bv1G009420	Expansin-B6	6B
GY	TRITD3Bv1G005570, TRITD3Bv1G006100, TRITD3Bv1G006040, TRITD3Bv1G005730, TRITD3Bv1G005840, TRITD3Bv1G006060, TRITD3Bv1G005850, TRITD3Bv1G008290	Glycoslytransferases	3B
TKW	TRITD2Av1G022960, TRITD2Av1G021680, and TRITD2Av1G021850	F-box protein	2A
DM	TRITD1Av1G141840, TRITD2Bv1G076800, TRITD5Av1G009310, and TRITD5Bv1G253520	F-box family protein	1A, 2B, 5A, 5B
SpD	TRITD1Av1G200090	F-box family protein	1A
P SDC, ZI	TRITD3Av1G251020, TRITD3Av1G251030, and TRITD3Av1G247490	F-box family protein	3A
DB	TRITD6Bv1G011600, TRITD6Bv1G011470, TRITD6Bv1G011220,TRITD6Bv1G011930, TRITD6Bv1G011950, TRITD6Bv1G009670, TRITD6Bv1G050250, and TRITD6Bv1G011630	F-box family protein	6B
VITP	TRITD5Bv1G167420, TRITD5Bv1G168190, and TRITD5Bv1G167310	F-box family protein	5B
GY	TdMYB6B213	Myb TF	6B
TKW	TdMYB3A049 and TdMYB3A050	Myb TF	3A
DB	TRITD1Av1G141030 and TRITD1Av1G140780, TdMYB5B202	Myb TF	1A,6B
PH	TRITD4Bv1G109310	Myb TF	4B
SDpSp	TRITD2Bv1G036410 and TRITD2Bv1G036450	Myb TF	2B
PSDC	TdMYB3A055	Myb TF	3A
GY	TRITD3Bv1G007750, TRITD3Bv1G007660	Cyclin-dk F-2	3B
GY	TRITD6Bv1G188320	Cdk inhibitor	6B
PH	TRITD5Bv1G196080, TRITD4Bv1G110750	WRKY TF	4B, 5B
TKW	TRITD6Bv1G010420	CytochromeP450 sf protein	6B
PH	TRITD2Av1G035460, TRITD2Av1G280860, and TRITD2Av1G281150, TRITD2Bv1G211170	CytochromeP450 sf protein	2A, 2B
Aw	TRITD4Bv1G145590, TRITD4Bv1G145600	CytochromeP450 sf protein	4B
WGC	TRITD2Bv1G060220, TRITD2Bv1G058980, TRITD2Bv1G058990, TRITD2Bv1G058930	CytochromeP450 sf protein	2B
PSDC,RSDC	TRITD7Bv1G205910, TRITD7Bv1G205620, TRITD7Bv1G210590, and TRITD7Bv1G210580	CytochromeP450 sf protein	7B
SDpSp	TRITD2Bv1G036110	CytochromeP450 sf protein	2B
SpD	TRITD2Bv1G256620	CytochromeP450 sf protein	2B
VITP	TRITD5Bv1G166720 and TRITD5Bv1G166710	CytochromeP450 sf protein	5B
VITP, ZI	TRITD5Bv1G165280, TRITD3Av1G248840, TRITD3Av1G248730, TRITD3Av1G248750	Glutathione S-transferase	5B

Remark: fs = super family.