

G OPEN ACCESS

Citation: Negisho K, Shibru S, Pillen K, Ordon F, Wehner G (2021) Genetic diversity of Ethiopian durum wheat landraces. PLoS ONE 16(2): e0247016. https://doi.org/10.1371/journal. pone.0247016

Editor: Manoj Prasad, National Institute of Plant Genome Research, INDIA

Received: August 6, 2020

Accepted: January 30, 2021

Published: February 17, 2021

Copyright: © 2021 Negisho et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its <u>Supporting</u> information files.

Funding: The research was funded by the Federal Ministry of Food and Agriculture (Bundesministerium für Ernährung und Landwirtschaft (BMEL), FKZ 2813FS01), Germany. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

RESEARCH ARTICLE

Genetic diversity of Ethiopian durum wheat landraces

Kefyalew Negisho¹, Surafel Shibru², Klaus Pillen³, Frank Ordon⁴, Gwendolin Wehner⁴*

1 Ethiopian Institute of Agricultural Research (EIAR), National Agricultural Biotechnology Research Center, Holeta, Ethiopia, 2 Ethiopian Institute of Agricultural Research (EIAR), Melkassa Research Center, Melkassa, Ethiopia, 3 Martin-Luther-University, Institute of Agricultural and Nutritional Sciences, Halle (Saale), Germany, 4 Julius Kühn Institute (JKI), Institute for Resistance Research and Stress Tolerance, Quedlinburg, Germany

* gwendolin.wehner@julius-kuehn.de

Abstract

Genetic diversity and population structure assessment in crops is essential for marker trait association, marker assisted breeding and crop germplasm conservation. We analyzed a set of 285 durum wheat accessions comprising 215 Ethiopian durum wheat landraces, 10 released durum wheat varieties, 10 advanced durum wheat lines from Ethiopia, and 50 durum wheat lines from CIMMYT. We investigated the genetic diversity and population structure for the complete panel as well as for the 215 landraces, separately based on 11,919 SNP markers with known physical positions. The whole panel was clustered into two populations representing on the one hand mainly the landraces, and on the other hand mainly released, advanced and CIMMYT lines. Further population structure analysis of the landraces uncovered 4 subgroups emphasizing the high degree of genetic diversity within Ethiopian durum landraces. Population structure based AMOVA for both sets unveiled significant (P < 0.001) variation between populations and within populations. Total variation within population accessions (81%, 76%) was higher than total variation between populations (19%, 24%) for both sets. Population structure analysis based genetic differentiation (FST) and gene flow (Nm) for the whole set and the Ethiopian landraces were 0.19 and 0.24, 1.04, and 0.81, respectively indicating high genetic differentiation and limited gene flow. Diversity indices verify that the landrace panel was more diverse with (I = 0.7, He =0.46, uHe = 0.46) than the advanced lines (I = 0.6, He = 0.42, uHe = 0.42). Similarly, differences within the landrace clusters were observed. In summary a high genetic diversity within Ethiopian durum wheat landraces was detected, which may be a target for national and international wheat improvement programs to exploit valuable traits for biotic and abiotic stresses.

Introduction

Durum wheat [*Triticum turgidum* ssp. *durum* (Desf.) Husn.] was domesticated from wild emmer (*Triticum turgidum* ssp. *dicoccoides*) to emmer (*Triticum turgidum* ssp. *dicoccum*)

followed by a secondary domestication, i.e. from emmer to naked forms and durum wheat (*Triticum turgidum* ssp. *durum*) [1]. The allotetraploidization event was involved after a cross between the two diploid species: *T. urartu* (genome AA) [2, 3] and an unknown close relative of *Aegilops speltoides* (genome BB) [4, 5]. Thus, durum wheat has an allotetraploid genome (AABB genome, 2n = 4x = 28, seven homoeologous groups—13,000 M bp) [6]. A high-density gene-associated SNP array was developed for the characterization of polyploid wheat [7] and complemented with fully annotated high-confident genes [8]. Maccaferri et al. [9] developed the high-density tetraploid wheat consensus map from data sets of durum wheat cultivars (*Triticum turgidum* ssp. *durum*), cultivated emmer (*T. turgidum* ssp. *dicoccum*) and their ancestor (wild emmer, *T. turgidum* ssp. *dicoccoides*). Recently, the reference sequence of the genome of cv. Svevo led to the identification of 66,559 high confidence (HC) genes enabling genome-wide genetic diversity analyses in tetraploid durum wheat [10].

Durum wheat is one of the ten most important crops worldwide with an annual production of 37 million tons [11] and Ethiopia is the major durum wheat producer in sub-Saharan Africa (SSA), with a durum acreage of 0.6 million ha [11–13]. Durum wheat is primarily used for pasta production, but in addition it is used to make flour for leavened biscuits, cookies, biofuel, and for fermentation to make alcoholic beverages such as beer and liquors [14]. In the country, durum wheat nearly accounts for 15–20% of wheat production and 30% of the whole acreage [15, 16]. Hence, it contributes about 18 to 20% to the national wheat production [17, 18]. In Ethiopia, wheat (both bread and durum) is produced by around 4.62 million households with an estimated land area of 1.7 million ha and mean national yield of 2.7 t/ha [19]. Traditionally, in Ethiopia wheat straw is used as animal feed and as roof thatching material. This makes wheat biomass highly valuable in rural communities. Thus, on top of high grain yield and environmental tolerance, in wheat growing areas farmers also take into account those traits when selecting landraces.

The Ethiopian Biodivesity Institute (EBI) hosts more than 7000 landraces collected from durum wheat growers for genetic conservation and for the exploitation of genetic diversity [20, 21]. Based on the genetic diversity analysis, Mengistu et al. [22] reported a high genetic variability in Ethiopian durum wheat landraces. Kabbaj et al. [11] have demonstrated that Ethiopian durum wheat landraces cluster separatly from durum of the International Center for Agricultural Research in the Dry Areas (ICARDA), Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), and durum wheat derived from other countries. Genetic diversity can be described as the degree of differentiation between or within species. Existing intra-and inter-specific differences are the base of all crop improvement programs [23]. Hence, genetic variation is an essential source of novel and useful alleles to be selected by breeders for abiotic and biotic resistance/tolerance [24, 25]. It is supposed that allelic variation of genes originally found in wild species, is gradually lost through domestication and breeding [26]. Therefore, the narrowed or lost allelic variation can be recovered by exploring e.g. landraces [26]. Landraces are genetically dynamic and are in equilibrium with biotic and abiotic stresses in the environments where they evolved [27, 28]. Therefore, landraces which have adapted to their natural environment over time [29-31] can contribute favorable genomic regions for tolerance against abiotic stresses like drought.

Analysis of genetic diversity in populations is an important topic in breeding as well as conservation and evolutionary genetics studies [32, 33]. Expected heterozygosity (He) or the genetic diversity index, which is derived from gene frequency data, is used to determine the genetic variation within populations [34, 35]. Wright [36], used the fixation index (FST) to estimate genetic differentiation among populations. Leinonen et al. [37] reported that FST estimated from DNA markers provide a starting point to assess the strength of divergent selection on quantitative traits. Gene flow (Nm), which is estimated through FST is used to estimate the gene exchange within population and among populations [38]. Additionally, genetic diversity indices provide useful information on genetic diversity. Genetic analyses, such as estimation of genetic diversity and population structure, as well as genome wide association studies and marker assisted selection procedures are broadly undertaken by molecular markers [39]. Single nucleotide polymorphisms (SNPs) and simple sequence repeats (SSRs) are the most common molecular markers in genetic studies [40, 41]. Out of these, SNP markers provide an increasing resolution due to their high abundance [25, 42, 43]. Additionally, the power of SNP markers in wheat recently elevated 100-fold from 9 K [44] to 820 K [45]. In this study, we used a hybridization array that includes about 90K SNPs, which was developed for genetic analyses in allohexaploid and allotetraploid wheat populations [7, 10].

Up to now only a small part of the huge collection of durum wheat landraces hosted at EBI was characterized using SSR [46–48] and SNP markers [22, 49]. Therefore, the aim of our study was to assess population structure and genetic diversity of 215 Ethiopia durum wheat landraces, 10 released durum wheat varieties, 10 advanced durum wheat lines from Ethiopia, and 50 CIMMYT durum wheat lines using highly informative SNP markers.

Material and methods

Plant material

A total of 285 durum wheat accessions, hereafter designated as study panel (SP) was used for the analysis of genetic diversity. The SP included, 215 Ethiopian durum wheat landraces assigned as ETDWL, 10 released durum wheat varieties, 10 advanced durum wheat lines from Ethiopia and 50 durum wheat lines from CIMMYT (S1 Table). The ETDWL were obtained from the Ethiopian Biodiversity Institute (EBI, http://www.ebi.gov.et/). Landraces were selected based on the acreage in each seed source region (origin). Thus, more samples were taken from major growing regions (Oromia and Amhara) and some samples from minor growing regions. 105 ETDWL were sampled from Oromia, 88 from Amhara, 1 from Benishangul Gumuz, 16 from Tigray and 5 from South Nation Nationalities and Peoples (SNNP), representing different seed sources (origin), seed collection zones and geographic regions (S1 Table). Online ArcGIS software was used to map the landraces collection areas in Ethiopia, https://www.arcgis.com/home/webmap/viewer.html, released version 10.8.1 July 2020 (Fig 1). For the Ethiopian durum wheat landraces, GPS passport data were obtained from EBI and are provided in S2 Table. A self-created layer was used to map positional data.

SNP genotyping

The durum wheat SP was grown in the greenhouse at Quedlinburg, Germany for 15 days under standard growing conditions, i.e. 20 to 22°C during day time and 17 to 19°C at night [50] with automatic water supply. Genomic DNA was extracted from single plant fresh leaves following the mini-prep DNA extraction protocol [51]. Genomic DNA quality was checked by 1% gel electrophoresis and DNA concentration measurement was conducted by NanoDrop[®] ND-1000 Spectrophotometer (Saveen Warner, Sweden). 50 ng of DNA per sample was used for SNP analysis using the 90K iSelect chip (Illumina Inc., San Diego, USA). Genotyping was conducted by Trait Genetics, Gatersleben (Germany). SNPs with a low minor allele frequency (MAF) are generally considered as rare with less power in detecting marker trait associations (MTAs) and are prone to genotyping error [52]. Thus, SNPs with minor allele frequency (MAF) of < 5%, missing data > 10% and heterozygosity > 12.5% were excluded from further analyses. Additionally, imputation was conducted using the software Beagle [53]. Physical SNP positions were taken from the reference sequence of durum wheat [8] to construct a hapmap file for further analyses.



Fig 1. Durum wheat landraces collection areas in Ethiopia. GPS: Geographic position system, Regions of seed origin: South Nation Nationalities and Peoples (SNNP, Light green), Benishangul Gumuz (red), Oromia (green), Amhara (pink), Tigray (blue), Ethiopian boundary and geo-positions were indicated. The mapping was performed using the online ArcGIS software suite vs. 10.8.1.

https://doi.org/10.1371/journal.pone.0247016.g001

Population structure and genetic diversity analyses

Genotypic data were used to describe the genetic diversity within the durum wheat study panel. We analyzed the population structure and genetic diversity of the ETDWL separately, and compared this with the population structure and genetic diversity of the SP. The underlying genetic population structure was estimated with STRUCTURE 2.3.4 software [54]. SNP markers having high polymorphic information content (PIC ≥ 0.35) were selected across all durum wheat chromosomes (S3 Table). Hence, 10,000 burn-in periods followed by 100,000 Markov Chain Monte Carlo (MCMC) iterations for K = 1 to 10 clusters were used to identify the optimal range of K with five replications per cluster for the SP as well as for ETDWL. The optimal K-value was determined using the Δ K method [55]. DARwin 6.0.17 [56] was used for molecular diversity analysis to get information on genetic dissimilarity among populations and within populations. Neighbor-Joining (NJ) algorithm of the genetic distances was determined according to Saitou and Nei [57] and used to create a phylogenetic tree.

Analysis of molecular variance (AMOVA) and genetic diversity indices

Genetic distance between populations was determined using Nei's Genetic Distance [58] based on the number of populations k. We run AMOVA, which allowed hierarchical partitioning of genetic diversity among populations and within populations [59]. Thus, AMOVA was

performed using GeneAlEx 6.503 [35]. Additionally, 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 [38]. Gene flow (Nm) among populations was calculated based on FST as:

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

In addition, Shannon's Information Index (I) [60], expected heterozygosity (He), unbiased heterozygosity (uHe), and the percentage of polymorphic loci (PPL) were calculated as follows:

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

 $He = 1 - \sum Pi^{2}, uHe = \left[\frac{N}{N-1}\right] * He$

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

Results

SNP analyses

After filtering, 11,919 SNPs were used for genetic analysis. These were continuously distributed across the A and B genome of durum wheat for the SP (Fig 2). In all cases the B genome showed a higher number of SNPs except for chromosome 7, for which 951 SNPs were detected on chromosome 7A and 911SNPs on chromosome 7B. The lowest number of SNPs were detected on chromosome 4A (553) and the highest SNP number was obtained for chromosome 2B (1237). Generally, in the current study, 58% of the SNPs were located on the B genome and 42% on the A genome.



Fig 2. Distribution of 11,919 filtered Single nucleotide polymorphisms (SNPs) across the durum wheat genome. Genome A and B are marked with red and blue colors, respectively.

https://doi.org/10.1371/journal.pone.0247016.g002

Population structure

Population structure analysis for the durum wheat SP revealed ΔK at K = 2 i.e two populations hereafter considered as Pop1 and Pop2 (Fig 3A–3C, S1 Table). Pop1 comprised 207 accessions. Of these, 206 accessions were from ETDWL and 1 from the durum wheat lines of Ethiopia. In relation to seed origin, the number of accessions in Pop1 were originated from Amhara (88), Oromia (97), SNNP(5), Tigray(16), and Debre Zeit Agricultural research Center (DZAR, 1) (Fig 3C). Pop2 constituted of 78 accessions. Fifty of the accessions in Pop2 were from CIM-MYT, 19 from the group of the released and advanced durum wheat lines of Ethiopia and 9 were landraces. The landraces clustered in Pop2 were DW006, DW007, DW008, DW020, DW039, DW143, DW185 and DW188 from Oromia, as well as DW050 from Amhara (Fig 3C, S1 Table). Thus, the SP mainly split in the ETDWL and advanced varieties.

Ethiopian durum wheat landraces (ETDWL) comprise accessions collected from major wheat producing regions of the country (Oromia, Amhara, Tigray and SNNP) (Fig 1, S1 Table). Population structure analysis of the ETDWL uncovers populations with Δ K at K = 4 (Fig 4A–4C, S1 Table). The populations in ETDWL comprised 45, 27, 47, and 96 accessions, respectively for Pop1-1, Pop1-2, Pop1-3, and Pop1-4 (Fig 4C). Pop1-1 comprises 45 accessions of which 22 originated from Amhara, 17 from Oromia, 2 from SNNP and 4 from Tigray. In the second cluster (Pop1-2) which comprises 27 accessions, 12 accessions were from Amhara, 14 from Oromia and 1 from SNNP. Pop1-3 consisted of 47 accessions, i.e. 14 from Amhara, 29 from Oromia, 1 from SNNP and 3 from Tigray. Pop1-4 comprises 96 accessions of which 41 derived from Amhara, 45 from Oromia, 1 from SNNP and 9 from Tigray. The number of



Fig 3. Population structure analysis of the durum wheat SP. (A) Bayesian structure analysis, (B) Structure harvester Evanno's test ΔK at K = 2. (C) Neighbor Joining (NJ). Populations identified in STRUCTURE based on Bayesian structure analysis are shown in red and green for Pop1 and Pop2, respectively.

https://doi.org/10.1371/journal.pone.0247016.g003





https://doi.org/10.1371/journal.pone.0247016.g004

accessions per cluster in ETDWL ranged from a minimum of 27 to a maximum of 96 for Pop1-2 and Pop1-4, respectively.

Analysis of molecular variance (AMOVA)

Analysis of molecular variance (AMOVA) for the durum wheat SP and ETDWL was conducted taking the respective population structure clusters (Δ K) into account (Table 1). In both cases, AMOVA indicated significant (P < 0.001) effects for variation between populations and within populations. The AMOVA of the SP revealed that 19% of the total variation is between populations, while 81% of the total variation is present within populations. Fixation index (FST) and gene flow (Nm) for the SP were calculated at FST = 0.19 and Nm = 1.04, respectively. Similarly, AMOVA for ETDWL revealed 24% of the total variation between populations and 76% variation within populations. Fixation index (FST) and gene flow (Nm) for the ETDWL were FST = 0.24 and Nm = 0.81, respectively. Therefore, the AMOVA for SP and ETDWL showed higher percentage of variation within populations than between populations (Table 1).

Source of variation	df	Sum of squares	Variance components	Percentage of variation	Fixation index (Fst)	Gene flow (Nm)
Variance partition of the	SP, $k = 2$					
Between populations	1	5323.9***	22.7	19	0.19	1.04
Within populations	283	53223.7***	94.0	81		
			Variance partition of th	e ETDWL, k = 4		
Between Populations	3	7069.6***	23.0	24	0.24	0.81
Within populations	211	31572.1***	74.8	76		

Table 1. AMOVA for the SP and ETDWL based on structure analysis results.

df: degree of freedom,

***: P-value at P < 0.001,

SP: Study panel, ETDWL: Ethiopian durum wheat landraces.

https://doi.org/10.1371/journal.pone.0247016.t001

Table 2. Mean of different genetic indices parameters in each population.

Рор	N	Ι	He	uHe	PPL					
Population of the SP, K = 2										
Pop1	207	0.7	0.46	0.46	100					
Pop2	78	0.61	0.42	0.42	100					
Population of	the ETDWL, k = 4									
Pop1-1	45	0.5	0.32	0.32	93.1					
Pop1-2	27	0.52	0.33	0.34	89.8					
Pop1-3	46	0.62	0.39	0.39	100					
Pop1-4	96	0.5	0.32	0.32	97.6					

Number of accessions (N), Shannon's information index (I): $I = -1 * \sum [Pi * LnPi]$, expected heterozygosity (He) or genetic diversity: $He = 1 - \sum Pi^2$, Unbiased heterozygosity $uHe = \left[\frac{N}{N-1}\right] * He$, and percentage of polymorphic loci: $PPL = \sum Pi/N$; SP: Study panel, ETDWL: Ethiopian durum wheat landraces.

https://doi.org/10.1371/journal.pone.0247016.t002

Genetic indices

We investigated the genetic diversity of the SP and ETDWL based on population structure analysis results of ΔK at k = 2 and k = 4, respectively (Table 2). The genetic indices for the SP such as I, He, and uHe showed higher values for Pop1 as compared to Pop2. Hence, Pop1 that comprised 99.5% accessions from ETDWL was more diverse (I = 0.7, He = 0.46, uHe = 0.46) than Pop2 (I = 0.6, He = 0.42, uHe = 0.42), which comprised 88.5% of improved varieties (advanced, released and CIMMYT durum wheat). Pop1-3 of the ETDWL was the most diverse (I = 0.62, He = 0.39, uHe = 0.39) with 100% PPL followed by Pop1-2 (I = 0.52, He = 0.33, uHe = 0.34) with 89.8% PPL. Pop1-1 and Pop1-4 showed similar genetic diversities (I = 0.5, He = 0.32, uHe = 0.32) with 93.1% and 97.6% PPL, respectively (Table 2).

Discussion

Hybridization arrays are believed to represent a significant fraction of SNPs distributed across genomes. In wheat they represent SNPs between populations of diverse geographical origin [7, 44, 45]. Hence, in this study, we used the hybridization array that includes about 90K SNPs, which was developed to analyze genetic variation in allohexaploid and allotetraploid wheat populations [7, 10]. Studies indicated a higher number of SNPs in the B than in the A genome of wheat [61, 62]. Likewise, higher number of SNPs was also identified in this study on the B

genome (58%) than on the A genome (42%) (Fig 2). However, we detected a higher number of SNPs on chromosome 7A (951) than on chromosome 7B (911). Similarly, studies by Naz et al. [63] and Desta et al. [64] on bread wheat indicated highest numbers of SNP markers on the B genome followed by the A genome and less across the D genome.

In this study, population structure and Neighbor-Joining (NJ) analysis showed two populations (Pop1 and Pop2) for the study panel (SP). Concerning Pop1, 206 (99.5%) accessions were from ETDWL and only 1, DZ005 (0.5%) from the advance durum wheat lines (Fig 3). This durum wheat line most probably was selected from landraces by Ethiopian durum wheat breeders. This elucidates that only little effort was spent to include landraces in durum wheat improvement programs in the country. Pop2 of the SP constituted of 69 (88.5%) accessions from CIMMYT and others that originate from international sources like ICARDA which are released durum wheat varieties and advanced durum wheat lines. The remaining 9 (11.5%) accessions are landraces. The landraces clustered in Pop2 were most probably incorrect renamings of the released durum wheat varieties as landraces during germplasm collection or they may be an admixture. In Ethiopian since 1970 until recently, CIMMYT is the major source for most of the improved durum wheat materials [23]. In support of this, this study clearly showed that 19 out of 20 Ethiopian accessions plus advanced durum lines are clustered in Pop2 with durum wheat lines from CIMMYT. This is possible under the scenario that most improved durum wheat materials were introduced from international breeding programs to the country [13]. Additionally, it shows that only little attention was given to explore the genetic diversity in ETDWL as pointed out by [22]. Therefore, in Ethiopia to exploit the existing genetic diversity more focus should be given to conserve and use the landraces in durum wheat breeding programs.

It has been reported that, Ethiopian durum wheat landraces are distinct and have no kinship with the Middle-East, which is the primary region of origin of durum wheat [11]. Therefore, the separate clustering of Ethiopian durum landraces from international varieties may illustrate a long-time separation of Ethiopian durum wheat landraces from primary durum origin and from international germplasm. This is attributed to the uniqueness of Ethiopian durum wheat landraces [11, 13, 22, 46]. This is in agreement with reports that designated Ethiopian durum wheat landraces as separate sub-species under the name *T. durum subs. abyssinicum or T. aethiopicum* [22, 23]. Additionally, separate clustering of Ethiopian durum wheat from improved durum wheat in Ethiopia indicated that little or no improved varieties were generated from landraces either through selection or via crossing with international durum wheat materials. Nevertheless, germplasm originating from international organization such as CIMMYT and ICARDA remain the main source for advanced durum lines and released durum varieties in Ethiopia [13].

Population structure analysis of ETDWL alone uncovers four populations (Pop1-1, Pop1-2, Pop1-3 and Pop1-4), which is in agreement with and NJ analysis (Fig 4). Mengistu et al. [22] have identified a number of populations (k = 10) in Ethiopian durum wheat landraces by removing improved durum wheat varieties from the population analysis. Our study also signifies the presence of higher admixture of accessions between different populations of landraces (Fig 4). This is a common phenomenon for most cereal crops grown in Ethiopia because of informal seed exchange systems involving regional and countrywide farming communities. In Ethiopia, farmers exchange seeds of cereals in various traditional forms such as gifts, barter, labor exchange or social obligations [65, 66]. Therefore, the main source of seed for planting wheat and barley landraces in Ethiopian smallholder communities is via the informal farmer to farmer seed exchange. Apparently, once farmers obtain seed with required quality that genotype will get bigger chance to spread across local communities. This was demonstrated by genetic clustering based on seed collection regions where seeds originated from one region

relatively closely clustered in the same population. For instance, 9 out of 16 accessions collected from Tigray were clustered in Pop1-4, on the other hand no accession from this region was grouped in Pop1-3 (Fig 4). The geographic isolation and latitudinal variation, i.e. 1540– 3190 meter above sea level from which accessions were collected, confirmed the variability and genetic dynamics in Ethiopian durum wheat landraces to adapt to wide-ranging conditions (Fig 1, S1 Table). Subsequently, the high-level of genetic diversity can be exploited in wheat breeding and improvement programs to overcome the biotic and abiotic stresses across latitudinal ranges.

Durum wheat is one of the important cereal crops grown in Ethiopia and the country is endowed with a wealth of genetic diversity for tetraploid wheat. Phenotypic and morphological analysis [23, 67–69] and genotypic analysis elucidated the existance of huge genetic diversity in ethiopian tetraploid wheat [22, 48, 70]. Consquently, the country is considered as the center of diversity and/or secondry center of origion for durum wheat [11, 22, 71]. In our study, genetic diversity within population accessions was higher than genetic diversity between populations (Table 1) illustrating that more attention should be given to individual accessions within populations to explore the existing genetic diversity as a basis for genomic analysis, and for genetic material conservation.

Fixation index (differentiation = FST) measures population differentiation due to genetic structure [72] and FST can be considered important in differentiating populations when its value is greater than 0.15 [73]. Hence, FST values were calculated at FST = 0.19 and FST = 0.24 for the SP and ETDWL, respectively indicating significant differentiations between the populations. Eventually, in our study, the higher genetic differentiation led to limited gene flow (Nm) values of Nm = 1.04 and Nm = 0.81 for the SP and ETDWL, respectively (Table 1). Nm value less than one is an indication of limited gene exchange as it was suggested by [38, 73]. Therefore, the Nm < 1 in ETDWL (0.81) clearly shows the high degree of genetic differentiation that exists among the ETDWL populations (FST = 0.24) as compared to SP (FST = 0.19) [37, 74]. In agreement to this, [39] reported that a high genetic exchange lead to low genetic differentiation between populations. Similarly [22], reported high genetic diversity in Ethiopian durum wheat landrace collections. Apparently, in the present study, population structure analysis of ETDWL alone revealed more populations suggesting the huge genetic diversity that exists within Ethiopian durum wheat landraces (Fig 4). Information on genetic diversity of each population can be assessed using genetic diversity indices [39]. Likwise, in this study, diversity analysis was further supported by the genetic diversity indices such as I, He, and uHe (Table 2). Genetic diversity indices for the SP illustrated higher genetic diversity in Pop1, which constituted 99.5% of the ETDWL as comparison to Pop2, which comprised only 11.5% ETDWL. Genetic diversity indices for the ETDWL indicated that Pop1-3 was the most diverse followed by Pop1-2, whereas Pop1-1 and Pop1-4 showed similar genetic diversity (Table 2). This marked that landraces showed huge genetic diversities that can broaden the genetic base for wheat improvement. In agreement to this, biotic and abiotic resistance/tolerance genes or genomic regions were identified in Ethiopian durum wheat landraces, e.g. resistance to biotic factors such as stripe rust resistance [75], adult plant resistance to leaf rust and stem rust [46, 76, 77], abiotic stress resistance such as aluminum tolerance [78] and terminal drought tolerance [23]. Therefore, Ethiopian durum wheat landraces may increase the rate of genetic gain if strategically included in wheat breeding programs. Most important, exploiting the landraces genetic diversity in Ethiopian durum wheat may help to mitigate abiotic stress factors that are apparent due to adverse effects of climate change. Furthermore, these landraces may help to uncover unknown genomic regions or genes associated with economically important traits.

Conclusion

We employed high quality SNP markers to analyze the population structure and genetic diversity of a durum wheat study panel comprising 285 accessions of which 215 accessions were ETDWL. AMOVA (P < 0.001) unveiled that genetic variation within population accessions was higher than genetic variation between populations for the SP and ETDWL. Structure analysis of SP revealed two distinct populations (Pop1 and Pop2). Genetic diversity indices for the SP illustrated higher genetic diversity in Pop1, which constituted 99.5% of the ETDWL in comparison to Pop2, which comprised only 11.5% ETDWL. Further population structure analysis of the ETDWL alone uncovered four populations emphasizing the high degree of genetic diversity that exists in ETDWL. Genetic diversity indices for the ETDWL indicated Pop1-3 was the most diverse followed by Pop1-2. Therefore, the high genetic diversity detected in ETDWL showed the existence of plentiful variability that could be utilized for future wheat breeding programs.

Supporting information

S1 Table. Population structure analysis results for SP. n = 285 with Delta K at k = 2 and ETDWL, n = 215 with Delta K at k = 4. (XLSX)

S2 Table. Passport data of the Ethiopian durum wheat landraces. (XLSX)

S3 Table. Highly informative selected 420 SNP markers for genetic diversity analysis and stratification.

(XLSX)

Acknowledgments

Our gratitude goes to the Ethiopian Biodiversity Institute (EBI), Debre Zeit Agricultural research Center (DZARC) and CIMMYT for sharing durum wheat accessions for the study panel. We are grateful to Julius Kühn Institute (JKI) technical, greenhouse and laboratory staff for their guidance and support. We would like to thank the technical staff of Ethiopian Institute of Agricultural Research (EIAR) at respective experimental sites for providing experimental plots and assistance.

Author Contributions

Conceptualization: Frank Ordon.

Data curation: Kefyalew Negisho, Surafel Shibru.

Formal analysis: Kefyalew Negisho.

Funding acquisition: Frank Ordon.

Methodology: Gwendolin Wehner.

Supervision: Klaus Pillen, Frank Ordon, Gwendolin Wehner.

Writing - original draft: Kefyalew Negisho.

Writing - review & editing: Klaus Pillen, Frank Ordon, Gwendolin Wehner.

References

- Gioia T, Nagel KA, Beleggia R, Fragasso M, Ficco DBM, Pieruschk R, et al. Impact of domestication on the phenotypic architecture of durum wheat under contrasting nitrogen fertilization. Journal of Experimental Botany. 2015; 66(18):5519–5530. https://doi.org/10.1093/jxb/erv289 PMID: 26071535
- 2. Konarev VG, Gavriluk IP, Peneva TI, Konarev AV, Khakimova AG, Migushova EF. On the Nature and origin of genomes of wheat based on the data of biochemistry and immunochemistry of grain proteins. Sel'skokhozyaistvennaya Biologia [Agricultural Biology]. 1976; 11(5):656–665.
- Dvorak JPE, McGuire, Cassidy B. Apparent sources of the A genomes of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. Genome. 1988; 30:680–689.
- Gill BS, Chen PD. Role of cytoplasm-specific introgression in the evolution of polyploid wheats. Proceedings of the National Academy of Sciences USA. 1987; 84:6800–6804. https://doi.org/10.1073/ pnas.84.19.6800 PMID: 16578821
- Kerby K, Kuspira J. The phylogeny of the polyploid wheats *Triticum aestivum* (bread wheat) and *Triticum durum* (macaroni wheat). Genome. 1987; 29:722–737.
- Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W. Genetics and geography of wild cereal domestication in the Near East. Nature Reviews Genetics. 2002; 3 (6):429–41. <u>https://doi.org/10.1038/</u> nrg817 PMID: 12042770
- Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, et al. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnology Journal. 2014; 12:787–796. https://doi.org/10.1111/pbi.12183 PMID: 24646323
- The International Wheat Genome Sequencing Consortium (IWGSC), Borrill. Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. Science. 2018; 361, eaar7191.
- Maccaferri M, Ricci A, Salvi S, Milner GS, et al. A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. Plant Biotechnology Journal. 2014; 13:648–663. https://doi.org/10.1111/pbi.12288 PMID: 25424506
- Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, et al, Durum wheat genome highlights past domestication signatures and future improvement targets. Nature Genetics. 2019; 51:885–95. https://doi.org/10.1038/s41588-019-0381-3 PMID: 30962619
- Kabbaj H. Sall AT, Al-Abdallat A, Geleta M, Amri A, Filali-Maltouf A, Belkadi B, et al. Genetic diversity within a global panel of durum wheat (*Triticum durum*) landraces and modern germplasm reveals the history of alleles exchange. Frontier in Plant Science. 2017; 8:1277. <u>https://doi.org/10.3389/fpls.2017</u>. 01277 PMID: 28769970
- Food and Agriculture Organization of the United Nations FAOSTAT statistics database, Crops. 2015; www.fao.org/faostat/en/#data/QC.
- **13.** Sall AT, Chiari T, Legesse W, Ahmed S, Ortiz R, van Ginkel M, et al. Durum wheat (*Triticum durum* Desf.) origin, cultivation, and potential expansion in sub-Saharan Africa. Agronomy. 2019; 9:263. https://doi.org/10.3390/agronomy9050263
- Tsegaye B, Berg T. Utilization of durum wheat landraces in East Shewa, central Ethiopia: Are home uses an incentive for on-farm conservation? Agriculture and Human Values. 2007; 24:219–230.
- Negassa A, Koo J, Sonder K, Shiferaw B, Smale M, Braun H J, et al. The potential for wheat production in sub–Saharan Africa: analysis of biophysical suitability and economic profitability. International Maize and Wheat Improvement Center (CIMMYT), Addis Ababa, Ethiopia; 2013.
- Alemu T, Zegeye H, Kassa D, Asnake D, Tafesse S, Asefa A. Wheat production concepts validation and assessment of dissemination and utilization constraints. Ethiopian institute of Agricultural Research. Research report No 126; 2019.
- 17. Tessema T, Bechere E. Developing elite durum wheat landrace selections (composites) for Ethiopian peasant farm use: raising productivity while keeping diversity alive. Euphytica. 1998; 323–328.
- Teklu Y, Hammer K. Diversity of Ethiopian tetraploid wheat germplasm: breeding opportunities for improving grain yield potential and quality traits. Plant Genetic Resource. 2008; 7:1–8.
- 19. CSA (Central statistical Agency of Ethiopia). Area and production of major crops, agricultural sample survey. Statistical bulletin 586. 2018; vol.1, April 2018, Addis Ababa, Ethiopia.
- IBC (Institute of Biodiversity Conservation). Ethiopia: Second Country Report on the State of PGRFA to FAO. 2013; Addis Ababa, Ethiopia. http://www.pgrfa.org.8.
- Mengistu DK, Kidane YG, Fadda C, Pè ME. Genetic diversity in Ethiopian durum wheat (*Triticum turgidum* var durum) inferred from phenotypic variations. Plant Genetic Resource. 2016; 16:39–49. https:// doi.org/10.1017/S1479262116000393

- Mengistu DK, Kidane YG, Catellani M, Frascaroli E, Fadda C, Pè ME, et al. High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. Plant Biotechnology Journal. 2016; 14(9):1800–1812. https://doi.org/ 10.1111/pbi.12538 PMID: 26853077
- Mengistu DK, Kiros AY, Pè ME. Phenotypic diversity in Ethiopian durum wheat (*Triticum turgidum* var. durum) landraces. Crop Journal. 2015; 3:190–199. https://doi.org/10.1016/j.cj.2015.04.003
- Bhandari HR, Bhanu AN, Srivastava K, Singh MN, Shreya, Hemantaranjan A. Assessment of Genetic Diversity in Crop Plants—An Overview. Advances in Plants and Agricultural Research. 2017; 7(3). https://doi.org/10.15406/apar.2017.07.00255
- 25. Acosta-Gallegos JA, Kelly JD, Gepts P. Pre-breeding and genetic diversity in common bean (*Phaseolus vulgaris*). Crop Science. 2007; 47:44–59.
- Fu YB. Understanding crop genetic diversity under modern plant breeding. Theoretical and Applied Genetics. 2015; 128:2131–2142. https://doi.org/10.1007/s00122-015-2585-y PMID: 26246331
- Lopes SM, El-Basyoni I, Baenziger PS, Singh S, Royo C, Ozbek K, et al. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. Journal of Experimental Botany. 2015; 66 (12):3477–3486. https://doi.org/10.1093/jxb/erv122 PMID: 25821073
- Mohammadi R, Sadeghzadeh B, Ahmadi H, Bahrami H, Amri A. Field evaluation of durum wheat landraces for prevailing abiotic and biotic stresses in highland rainfed regions of Iran. The Crop Journal. 2015; 3(5):423–433.
- Reynolds M, Dreccer F, Trethowan R. Drought adaptive traits derived from wheat wild relatives and landraces. Journal of Experimental Botany. 2007; 58(2):177–186. <u>https://doi.org/10.1093/jxb/erl250</u> PMID: 17185737
- 30. Acquaah G. Principles of plant genetics and breeding. 2nd ed. John Wiley and Sons, Ltd; 2012.
- **31.** Brown AHD. The genetic structure of crop landraces and the challenge to conserve them in situ on farms. In: Brush SB, editor. Genes in the field: on-farm conservation of crop diversity. Boca Raton: Lewis Publishers; 2000. 29–48.
- Jost L, Archer F, Flanagan S, Gaggitti O, Hoban S, Latch E. Differentiation measures for conservation genetics. Evolutionary Applications. 2018; 11(7). https://doi.org/10.1111/eva.12590 PMID: 30026802
- **33.** Caballero A, Toro MA. Analysis of genetic diversity for the management of conserved subdivided populations. Conservation Genetics. 2002; 3:289–299.
- Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes. 2006; 6(1):288–295.
- Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 2012; 28(19):2537–2539. <u>https://doi.org/10.1093/</u> bioinformatics/bts460 PMID: 22820204
- Wright S. Evolution and the Genetics of Populations: The Theory of Gene Frequencies. The University of Chicago Press, Chicago, Illinois; 1969.
- Leinonen T, O'hara R, Cano J Merila J. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. Journal of Evolutionary Biology. 2008; 21(1):1–17. https://doi.org/10. 1111/j.1420-9101.2007.01445.x PMID: 18028355
- Hartl DL, Clark AG. Principles of Population Genetics 3rd Ed. Sunderland, Massachusetts: Sinauer Associates, Inc; 1997.
- Eltaher S, Sallam A, Belamkar V, Emara HA, Nower AA, Salem KFM, et al. Genetic diversity and population structure of F3:6 Nebraska winter wheat genotypes using genotyping-by-sequencing. Frontiers in Genetics. 2018; 9. https://doi.org/10.3389/fgene.2018.00076 PMID: 29593779
- **40.** Rafalski A. Applications of single nucleotide polymorphism in crop genetics. Current Opinion in Plant Biology. 2002; 5:94–100. https://doi.org/10.1016/s1369-5266(02)00240-6 PMID: 11856602
- Kumar S, Banks TW Cloutier S. SNP discovery through next generation sequencing and its applications. International Journal of Plant Genomics. 2012; 12. <u>https://doi.org/10.1155/2012/831460</u> PMID: 23227038
- Govindaraj M, Vetriventhan M, Srinivasan M. Review article importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics Research International. 2015;
- Hyten DL, Song Q, Zhu Y, Choi IY, Nelson RL, Costa JM, et al. Impacts of genetic bottlenecks on soybean genome diversity. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103:16666–16671. https://doi.org/10.1073/pnas.0604379103 PMID: 17068128
- 44. Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, et al. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars

Proceedings of the National Academy of Sciences of the United States of America. 2013; 110 (20):8057–62. https://doi.org/10.1073/pnas.1217133110 PMID: 23630259

- 45. Winfield MO, Allen AM, Burridge AJ, Barker GLA, Benbow HR, Wilkinson PA, et al. High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. Plant Biotechnology Journal. 2016; 14:1195–206. https://doi.org/10.1111/pbi.12485 PMID: 26466852
- 46. Haile JK, Hammer K, Badebo A, Nachit MM, Röder MS. 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. 2013; 60:513–27. <u>https://doi.org/10. 1017/S1479262116000393</u>
- Asmamaw M, Keneni G, Tesfaye K. Genetic diversity of Ethiopian durum wheat (*Triticum durum* desf.) Landrace collections as reveled by ssr markers. Advances in Crop Science and Technology. 2019; 7 (1):413. https://doi.org/10.4172/2329-8863.1000413
- Teklu Y, Hammer K, Huang XQ, Roder MS. Analysis of microsatellite diversity in Ethiopian tetraploid wheat landraces. Genetetic Research and Crop Evolution. 2006; 53(6):1115–1126. <u>https://doi.org/10. 1007/s10722-005-1146-7</u>
- 49. Alemu A, Feyissa T, Letta T, Abeyo B. Genetic diversity and population structure analysis based on the high-density SNP markers in Ethiopian durum wheat (*Triticum turgidum* ssp. durum). BioMed Central Genetics. 2020; 21:18. https://doi.org/10.1186/s12863-020-0825-x PMID: 32050895
- Wehner G, Balko C, Ordon FF. Experimental design to determine drought stress response and early leaf senescence in barley (*Hordeum vulgare* L.). Bio-Protocol. 2016; 6(5).
- Stein N, Herren G, Keller B. A new DNA extraction method for high-throughput marker analysis in a large-genome species such as *Triticum aestivum*. Plant Breeding. 2008; <u>https://doi.org/10.1046/j.1439-0523.2001.00615.x</u>
- Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. The International Journal of Methods in Psychiatric Research. 2017; https://doi.org/10.1002/mpr.1608 PMID: 29484742
- Browning SR. Missing data imputation and haplotype phase inference for genome-wide association studies. Human Genetics. 2008; 124(5):439–50. https://doi.org/10.1007/s00439-008-0568-7 PMID: 18850115
- Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology. 2005; 14(8):2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x PMID: 15969739
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155:945–959. PMID: 10835412
- Perrier X, Flori A, Bonnot F. Data analysis methods. In: Hamon P., Seguin M., Perrier X., Glaszmann J. C. Ed., Genetic diversity of cultivated tropical plants. Enfield, Science Publishers. Montpellier; 2003. 43–76.
- Saitou N, Nei M. The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution.1987; 4(4): 406–425. https://doi.org/10.1093/oxfordjournals.molbev. a040454 PMID: 3447015
- Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 1978; 89:583–590. PMID: 17248844
- Meirmans PG, Liu S. Analysis of molecular variance (AMOVA) for autopolyploids. Frontiers in Ecology and Evolution. 2018; 6:66.
- Brown AHD, Weir BS. Measuring genetic variability in plant populations, in Isozymes in Plant Genetics and Breeding, Part A., Tanksley SD and Orton TJ, Editors. Elsevier Science Publishing. Amsterdam. 1983; 219–239.
- Alipour H, Bihamta MR, Mohammadi V, Peyghambari SA, Bai G, Zhang G. Genotyping-by-sequencing (gbs) revealed molecular genetic diversity of iranian wheat landraces and cultivars. Frontiers in Plant Science. 2017; 8:1293. https://doi.org/10.3389/fpls.2017.01293 PMID: 28912785
- Poland JA, Brown PJ, Sorrells ME, Jannink J. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping by- sequencing approach. PLOS ONE. 2012; 7(2). https://doi.org/10.1371/journal.pone.0032253 PMID: 22389690
- **63.** Naz AA, Dadshan S, Ballvora A, Pillen K, Léon J. Genetic analysis and transfer of favorable exotic QTL alleles for grain yield across d genome using two advanced backcross wheat populations. Frontiers in Plant Science. 2019; https://doi.org/10.3389/fpls.2019.00711 PMID: 31214227
- Desta ZA, Orabi J, Jahoor A, Backes G. Genetic diversity and structure found in samples of Eritrean bread wheat. Plant Genetic Resources. 2014; 12(01). https://doi.org/10.1017/S1479262113000324

- Bishaw Z. Wheat and barley seed systems in Ethiopia and Syria. PhD Thesis, Wageningen University, The Netherlands; 2004; 383. https://edepot.wur.nl/121548.
- Hailye A, Verkuijl H, Mwangi W, Yalew A. Farmers' wheat seed sources and seed management in the Enebssie area, Ethiopia. CIMMYT, Mexico and EARO, Ethiopia. 1998; 32.
- Eticha F, Bekele E, Belay G, Börner A. Phenotypic diversity in tetraploid wheats collected from Bale and Wello regions of Ethiopia. Plant Genetic Resources. 2005; 3(1):35–43. <u>https://doi.org/10.1079/</u> PGR200457
- Eticha F, Belay G, Bekele E. Species diversity in wheat landrace populations from two regions of Ethiopia. Genetic Resources and Crop Evolution. 2006; 53: 387–393. https://doi.org/10.1007/s10722-004-6095-z
- Geleta N, Grausgruber H. Morphological and quality traits variation in tetraploid (*Triticum turgiduml*.) and hexaploid (*Triticum aestivuml*.) wheat accessions from Ethiopia. Agricultural Science Research Journals.2013; 3(8):229–236.
- Badaeva ED, Shishkinaa AA, Goncharovc NP, Zuevd EV, Lysenkod NS, Mitrofanovad OP, et al. Evolution of *Triticum aethiopicum* jakubz. From the position of chromosome analysis. 2018; Genetika. 54(6):613–628. https://doi.org/10.1134/S1022795418060029
- Vavilov NI. The origin, variation, immunity and breeding of cultivated plants. 1951; Chronica Botanica. 13:1–351.
- 72. Luo Z, Brock J, Dyer JM, Kutchan T, Schachtman D, Augustin M, et al. Genetic diversity and population structure of a *Camelina sativa* spring panel. Frontiers in Plant Science. 2019; <u>https://doi.org/10.3389/</u> fpls.2019.00184 PMID: 30842785
- 73. Frankham R, et al., Introduction to Conservation Genetics. Cambridge University Press. Cambridge; 2002.
- 74. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution. 1965; 19:395–420.
- 75. Alemu SK, Badebo A, Tesfaye K, Uauy C. Identification of stripe rust resistance in Ethiopian durum wheat by phenotypic screening and Kompetitive Allele Specific PCR (KASP) SNP markers. Journal of Plant Pathology and Microbiology. 2019; 10(8). https://doi.org/10.35248/2157-7471.19.10.483
- 76. Habtamu A. Evaluation of ethiopian durum wheat varieties and landrace cultivars for the adult plant resistance of wheat leaf rust (*Puccinia triticina*). International Journal of Advances in Scientific Research and Engineering. 2019; https://doi.org/10.31695/IJASRE.33073
- Klindworth D, Miller J, Jin Y, Xu S. Chromosomal of genes for stem rust resistance in monogenic lines derived from tetraploid wheat accession ST464. Crop Science. 2007; 47:1012–1013.
- Wayima EF, Ligaba-Osena A, Dagne K, Tesfaye K, Machuka EM, Mutiga SK, et al. Screening of diverse Ethiopian durum wheat accessions for Aluminum tolerance. Agronomy Journal. 2019; 9(9):440. https://doi.org/10.3390/agronomy9080440