Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Morphological characterization and Genetic diversity analysis of a Tunisian durum wheat (*Triticum turgidum* var. *durum*) collection

Maroua Ouaja
Institut National Agronomique de Tunis
Bochra Amina Bahri
Institut National Agronomique de Tunis
Lamia Aouini
Institut National Agronomique de Tunis
Sahbi Ferjaoui
CRRGC
Maher Medini
BNG
Thierry Marcel
AgroParisTech
Sonia Mihed Hamza (hamza.sonia@inat.agrinet.tn)
Institut National Agronomique de Tunisie https://orcid.org/0000-0002-7036-143X

Research article

Keywords: Durum wheat, local landraces, landrace characterization, phenotypic diversity, genetic diversity, population structure

Posted Date: December 4th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-51248/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at BMC Genomic Data on February 3rd, 2021. See the published version at https://doi.org/10.1186/s12863-021-00958-3.

Abstract

Background: Tunisia is a center of genetic diversity of durum wheat and has a large number of abandoned old local landraces. An accurate investigation and characterization of the morphological and genetic features of these landraces would allow their rehabilitation and use for practical and beneficial purposes. In this context, a collection of 304 local accessions of durum wheat, collected from five regions and three climatic zones of central and southern Tunisia, was studied.

Results: Morphological characterization was carried out using 12 spike-related traits and rendered a mean Shannon-Weaver Index (*H'*) of 0.80 indicating the presence of a high level of polymorphism among accessions. Based on these traits 11 local landraces, namely Mahmoudi, Azizi, Jneh Khotifa, Mekki, Biskri, Taganrog, Biada, Badri, Richi, Roussia and Souri were identified. Spike length (*H'*=0.98) and shape (*H'*=0.86) with grains size (*H'*=0.94), form (*H'*=0.87) and color (*H'*=0.86) were the most polymorphic morphological traits. The genetic diversity was assessed using 10 SSR markers, with a polymorphic information content (PIC) of 0.69. Levels of genetic diversity were generally high, with a Shannon's Information Index (*I)* of 0.62 and a gene diversity (*He*) of 0.35. In addition, population structure analysis distinguished 11 genetic subpopulations significantly correlated with the morphological identification. Analysis of molecular variance (AMOVA) showed high genetic variations within regions (81%) and within wheat subpopulations (41%) reflecting a considerable amount of admixture between landraces. The moderate (19%) and high (59%) genetic variations among regions and among wheat subpopulations observed highlighted farmers selection practices . Furthermore, Mahmoudi landrace showed spike densities significantly different between the center to the south of Tunisia; notably loose spikes with open glumes in the south and compact ones in the center, which may represent an adaptation form for tolerance to high temperature.

Conclusion: Overall, this study underlined the genetic richness of local resources for better *in situ* or *ex situ* conservation and for their subsequent use in plant breeding programs.

Background

Durum wheat, *Triticum turgidum* var. *durum* Desf. (2n=4x=28, AABB), is a traditionally worldwide cultivated crop especially in the Mediterranean Basin. It originated from the Fertile Crescent and spread through different patterns of dispersal within the Mediterranean region, reaching the Iberian Peninsula through Northern Africa around 7000 BC. Since then, it has become a commercially important tetraploid wheat species with a center of diversification and production mainly centered in the Mediterranean Basin [1, 2]. This region is characterized by highly variable environments, from warm and dry to cool and wet climates [3]. Durum wheat germplasm collections from the Mediterranean area are characterized by a higher genetic diversity than the collections from other regions of the world [4].

Within the Mediterranean, Tunisia is one of the centers of diversity for durum wheat [5, 6]. Old Tunisian durum wheat cultivars are known by their high level of genetic diversity and their specific adaptability to North African drylands [7]. Despite their notable genetic diversity, Tunisian landraces were progressively abandoned since the first decade of the twentieth century and replaced by improved, high-yield and genetically uniform semi-dwarf cultivars (known as “modern varieties”) which were derived from international breeding programs [8,9]. This has led to a significant loss of the genetic diversity of the local durum wheat [10, 11]. Nonetheless, conserving the genetic diversity of durum wheat would be still possible by (1) the characterization of the remaining durum wheat landraces, (2) their re-introduction into breeding programs and (3) their protection through effective conservation strategies. Therefore, the genetic and morpho-phenological characterization of landraces, sparsely cultivated under current farming system or stored in gene banks, would allow the identification of unexplored sources of diversity that may contain adaptation to several biotic and abiotic stresses [4, 12, 13]. The availability of landraces for breeding programs may also have particular relevance when breeding for suboptimal and marginal environments such as the Mediterranean Basin, where durum wheat and other crop species are largely cultivated under unstable and limited water conditions that cause considerable yield fluctuations [14, 15].

Previous morphological characterization of old durum wheat germplasm from Tunisia, recorded 40 durum wheat landraces [10]. Agro-morphological evaluation of Tunisian durum wheat collections using quantitative and qualitative traits related to different parts of the spike, mostly grains, revealed high morphological diversity within the durum wheat landrace collection of Tunisia [16, 17]. However, few studies were conducted on the description of morphological and genetic features of durum wheat simultaneously. Moreover, the correlation between genetic population structure and morphological aspects in durum wheat was never investigated. The levels of genetic diversity of Tunisian durum wheat germplasm were assessed by Medini *et al.* [7] using AFLP and SSR markers which reveals an important polymorphism within cultivars. More recently, Robbana *et al.* [18] investigated the genetic diversity and population structure of 196 durum wheat landrace accessions (including Tunisian and North African accessions) using DArTseq markers. Their results showed that genetic variation was higher among landraces than within them, with a remarkable genetic similarity between the Tunisian and the North African landraces. Furthermore, Slim *et al.* [19] evaluated the genetic structure of Tunisian durum wheat germplasm and suggested the existence of five subpopulations with a strong genetic gradient from the north to the south of Tunisia, probably due to the prevalence of modern cultivars in the north. By tracing the history of cultivation, Tunisian durum wheat germplasm collections have been divided into three distinct categories; namely, traditional varieties or old landraces, old cultivars (cultivated up to 1970s), and modern cultivars (cultivated up to present) [7, 10, 19]. Traditional local landraces might harbor key traits for breeding programs, as they have been derived either from artificial selection of traditional farming systems or from a natural adaptation to adverse growing conditions.

Within this context, the objectives of the present study were (i) to evaluate the genetic diversity and population structure of Tunisian durum wheat accessions collected from central and southern Tunisia using SSR markers, (ii) to study the phenotypic diversity based on spike morphological characterization and (iii) to analyze the relationship between genetic and phenotypic variation.

Results

Morphological characterization of the Tunisian durum wheat accessions

Phenotypic diversity and morphological characterization

The Shannon-Weaver index (*H'*) revealed a high morphological diversity among the durum wheat accessions with an overall *H'* of 0.80 (**Table 1**). The most polymorphic characters were the spike length (*H'*=0.98), the grain size (*H'*=0.94), grain forms (*H'*=0.87), the grain color (*H'*=0.86) and the spike shape (*H'*=0.86), while the spike color showed the least polymorphic level (*H'*=0.53).

Table 1. Shannon-Weaver index (*H'*) estimated on the 304 Tunisian durum wheat accessions for the different regions and for the different climatic stages.

		Phenotypic traits												
		SS	SL	AL	SC	NS	GlC	GN	GF	GS	GC	AC	SD	Mean *H'*
Collection		0.86	0.98	0.79	0.53	0.83	0.84	0.69	0.87	0.94	0.86	0.64	0.74	**0.80**
Regions	Sousse	0	0	0	0	0	0	0	0	0	0	0	0	0.00
	Mahdia	0.52	0.99	0.56	0.00	0.63	0.48	0.62	0.63	0.48	0.74	0.73	0.65	0.58
	Kairouan	0.85	0.98	0.57	0.48	0.96	0.97	0.62	0.98	0.88	0.75	0.12	0.70	0.74
	Gabes	0.44	0.49	0.50	0.25	0.63	0.50	0.57	0.56	0.43	0.73	0.61	0.65	0.53
	Medenine	0.60	0.71	0.86	0.53	0.59	0.53	0.78	0.86	0.63	0.62	0.71	0.47	0.66
	Mean	0.54	0.69	0.55	0.30	0.61	0.55	0.55	0.65	0.56	0.62	0.47	0.53	**0.55**
Climatic stages	LSA	0.48	0.00	0.92	0.63	0.61	0.41	0.62	0.61	0.62	0.74	0.67	0.79	0.59
	MA	0.68	0.38	0.80	0.87	0.62	0.56	0.70	0.88	0.54	0.89	0.71	0.65	0.69
	HA	0.85	0.48	0.97	0.58	0.96	0.96	0.63	0.98	0.88	0.75	0.12	0.68	0.74
	Mean	0.67	0.29	0.90	0.69	0.73	0.64	0.65	0.82	0.68	0.79	0.50	0.71	**0.67**

SS : Spike Shape ; **SL :** Spike Length ; **AL :** Awns Length ; **SC :** Spike Color ; **NS :** Number of Spikelets/Spike ; **GlC :** Glumes Color ; **GN :** Number of Grains/Spikelet ; **GF :** Grains Form ; **GS :** Grains Size ; **GC :** Grains Color ; **AC :** Awns Color ; **SD :** Spike Density ; **LSA :** Low Semi-Arid (Sousse and Mahdia) ; **MA :** Mid-Arid (Gabes and Medenine) ; **HA :** Higher-Arid (Kairouan)

The 304 durum wheat accessions were grouped into eleven landraces namely Azizi, Jneh Khotifa, Taganrog, Mekki, Richi, Souri, Roussia, Badri, Biskri, Biada and Mahmoudi, according to the catalog of old durum wheat landraces and are part of the 40 durum wheat landraces recorded in Tunisia [10]. These landraces were characterized by the 12 specific morphological traits based on IPGRI [20] and UPOV [21] (**Table S1, Table S2)**. A multitude of spike characteristics has been observed between the durum wheat landraces, whereas these characteristics were homogeneous between accessions of the same landrace. In fact, the Shannon-Weaver index (*H'*) calculated for each landrace were relatively low, ranging between *H'*=0.00 for Badri and Jneh Khotifa landraces and *H'*=0.23 for Richi landrace with an overall mean *H'* of 0.11 (**Table S3**). For instance, the variety Mahmoudi accessions had particularly large spikes with sub-pyramidal shape, very long awns and big grain size; whereas rectangular and very flat spikes characterized Azizi accessions. Biskri accessions had fusiform and big size spikes. The Spike color, length and shape were variable among the studied accessions and varied from dark to light, and from short to long spikes For example, Badri spike was very short and thick with a greyish color, while Biada was characterized by very light (white) spikes and awns color. Souri and Roussia were both characterized by tight and red colored spikes with distinct spike shape characterized as rectangular for Souri and cylindrical for Roussia. The former varieties were also characterized by a distinct orange grain color. Interestingly, Richi accessions had a unique feathery spike, while Mekki was characterized by short and dense spikes with parallel edges. Finally, Taganrog accessions are characterized by white colored spikes washed with black, while Jneh Khotifa accessions had a very dark (black to purple) long and dense spike and awn colors.

Principle Coordinates Analysis (PCoA)

The PCoA performed on the 12 spike morphological traits of the 304 durum wheat accessions showed that axes 1 and 2 accounted for 25.73% and 22.34% of the total genetic variation, respectively (**Figure 1**). The first axis was mostly associated with Spike Shape (SS), Spike Length (SL), Number of Spikelets/spike (NS), Grain Color (GC) and Awns Length (AL). While the second axis was mainly defined by Grain Form (GF), Size (GS) and Number/spikelet (GN) (**Figure 1, a**). Color-coding of the accessions in the 2-dimensional PCoA plot (axis 1 vs. 2) showed a good correspondence between the morphological grouping and the landraces denomination (**Figure 1, b**). In fact, accessions belonging to the same landrace were included in the same PCoA subgroup. Biskri, Jneh Khotifa and Taganrog accessions were agglomerated and positively correlated to both axes and shared similar spike characteristics such as spike length (mostly long spikes) with a high number of grains per spikelet (>3) and similar black awns longer than the spike. Azizi accessions were grouped into a distinct subgroup mainly characterized by rectangular medium-sized spikes with a tan color. Mahmoudi accessions also formed a distinct subgroup that was mainly characterized by unique pyramid shaped spikes. Accessions of Souri and Roussia formed almost a single subgroup characterized by red colored loose and long spikes, as well as red colored glumes and awns. Landraces Badri and Mekki formed distinct subgroups negatively correlated to axes 1 and 2 and were both mainly characterized by short spikes with a low to intermediate number of grains per spikelet. Biada and Richi accessions were grouped mainly in the center of the plot and were particularly characterized by white colored spikes, glumes and awns (**Table S2**). Overall, all landraces were morphologically distinguished using the two axes based on the 12 spike characteristics, except for landraces Roussia and Souri and for landraces Biskri, Jneh Khotifa and Taganrog that were not distinct from each other in respect to their spike size and color. Thus, additional morphological traits were considered to classify the latter landraces into distinct subgroups such as glumes form (**Table S2**).

Genetic diversity and population structure of Tunisian durum wheat accessions

SSR polymorphism

Ten SSR markers were used in this study and were mapped on eight different chromosomes and considered therefore largely independent (**Table 2, Table S4**). The percentage of missing data was low and always remained below 10% for each locus. The 10 SSR markers amplified a total of 99 alleles. The number of different alleles per locus (Na) varied from 4 for Xgpw2103 to 16 for Xgwm413, with a mean of 9.9 across all loci. Overall, the PIC value was 0.690. The highest PIC value was obtained for Xgwm413 (0.851), whereas the lowest PIC value was obtained for Xgpw2103 (0.448). The Shannon's information index (*I*) also showed the highest value for Xgwm413 (2.182), whereas the lowest *I* value was obtained for Xgpw2103 (0.781). The fixation index (*Fis*) was close to 1 for each locus except for Xgwm495 (*Fis*= -0.373), where a high PIC level was observed (0.659). Pairwise genetic differentiation (F*st* value) ranged from 0.201 for Xgwm495 to 0.688 for Xgpw7148.

Table 2. Polymorphism level of the 10 Simple Sequence Repeats (SSR) markers used on 302 Tunisian durum wheat accessions.

Locus	N	Na	*I*	*Fis*	F*st*	PIC
Xgwm413	302	16	2.182	0.987	0.337	0.851
Xgpw7148	302	8	1.294	1.000	0.688	0.665
Xgwm495	300	11	1.614	-0.373	0.201	0.659
Xgwm193	298	10	1.338	1.000	0.577	0.621
Xgpw2239	302	8	1.695	1.000	0.424	0.773
Xgwm285	299	12	1.832	0.965	0.624	0.805
Xgpw4082	282	7	1.324	1.000	0.737	0.632
Xgpw4004	278	11	1.546	1.000	0.589	0.740
Xgpw2103	291	4	0.781	1.000	0.523	0.448
Xgwm372	275	12	1.643	0.988	0.491	0.705
Total	292.9 (3.378)	9.9 (1.048)	1.525 (0.118)	0.857 (0.137)	0.519 (0.052)	0.690

N : Samples size ; **Na :** Number of Alleles ; ***I* :** Shannon's Information Index ; ***Fis* :** Inbreeding coefficient within individuals ; **F*st* :** Inbreeding coefficient within subpopulations ; **PIC :** Polymorphic Information Content

Population structure analysis and relationship between genetic and morphological characterizations

A population structure analysis was investigated using the 302 Tunisian durum wheat accessions (188 MLG). The maximum likelihood (LnP(K)) and delta K (ΔK) methods indicated that the most likely number of genetic groups (K) was 11 (**Figure 2, a and b**). The estimated membership coefficients of each accession to the different genetic groups (at K=11) is shown in the population structure plot (**Figure 2, c)**.

Overall, each genetic group corresponds to a landrace. The genetic groups G2, G3, G4, G5, G7, G9, G10 and G11 corresponded to Jneh Khotifa, Taganrog, Mekki, Richi, Badri, Beskri, Biada and Mahmoudi, respectively. Moreover, a significant correlation between the genetic distance matrix and morphological distance matrix was obtained (P=0.01; R_{xy}=0.435). However, a discrepancy between the genetic distance matrix and the morphological distance matrix was observed for the landraces Azizi, Souri and Roussia. In fact, Azizi was clustered by STRUCTURE into two different genetic groups G1 and G8, and the two landraces Souri and Roussia were clustered in one genetic group G6 despite their distinct morphological characters.

Forty-one admixed individuals were observed in the collection. The majority of the admix is composed by G6 (Roussia and Souri) and G10 (Biada) representing 14.6 % of the admixed genotypes, followed by G1 (Azizi) and G9 (Beskri) representing 9.7 %.

Mahmoudi G11, Beskri G9 and the admixed accessions were the most frequent, composing the overall landrace collection, with 23.8 %, 12.2% and 14%, respectively. Azizi G1, Taganrog G3, Mekki G4, Badri G7 and Biada G10 each accounted for about 8% of the entire collection. However, Jneh Khotifa G2, Richi G5, Roussia and Souri G6 and Azizi G8 were the least represented in the collection and each accounted for solely 3% of the collection.

Analysis of diversity indices and molecular variance

The eleven clusters defined by the STRUCTURE analysis presented different levels of genetic diversity (**Table 3**). Group G6 showed the highest level of genetic diversity, while G7 represented the lowest level. The number of effective alleles per locus ranged from 1.152 for G7 to 2.379 for G6. Genetic groups with the highest number of MLGs were G6 (100% of different MLGs), G8 (90%) and G3 (85.7%), while G7 and G11 had the lowest number of MLGs, with 27.2% and 34.7%, respectively. The percentage of polymorphism ranged from 40% for G7 to 100% for G6 and G8. Shannon's index varied from 0.166 for G7 to 0.937 for G6 with an average of 0.620 across all accessions. In addition, G6 and G8 had the highest number of private alleles, with seven and four private alleles respectively; while G2 and G7 had no private alleles (**Table S5**). G10 and G4 had both two diagnostic alleles, while G3, G5 and G7 had one diagnostic allele with a frequency > 70%. The fixation index (*Fis*) ranged from 0.698 for G4 to 1 for G7 where *Ho* was 0.100 and null, respectively. Furthermore, the analysis of variance showed that 59% of the total genetic diversity was observed between the distinct genetic groups, while 41% of the genetic diversity was explained by differences within each group (**Table 4**).

Table 3. Diversity indexes of 302 Tunisian durum wheat accessions grouped by genetic subpopulations as defined by STRUCTURE, by regions and by climatic stages.

		Acc	MLG	S	Ne	*I*	*Ho*	*He*	*Fis*	P (%)	PA	*Nm*	Var-Pop	DA*
Subpopulations	ADMIX	41	33	13	3.904 (0.387)	1.522 (0.107)	0.088 (0.080)	0.721 (0.027)	0.871 (0.118)	100	6		-	-
	G1	24	17	5	1.830 (0.331)	0.627 (0.156)	0.033 (0.029)	0.334 (0.078)	0.869 (0.118)	90	3		100% Azizi	-
	G10	21	14	4	1.591 (0.229)	0.431 (0.144)	0.105 (0.100)	0.261 (0.088)	0.726 (0.201)	60	1		100% Biada	179 (Xgwm193) 214 (Xgpw4082)
	G11	72	25	11	1.443 (0.205)	0.394 (0.144)	0.099 (0.099)	0.210 (0.079)	0.784 (0.180)	70	1		100% Mahmoudi	-
	G2	9	6	3	1.694 (0.186)	0.510 (0.127)	0.111 (0.099)	0.332 (0.080)	0.688 (0.236)	70	0		100% JK	-
	G3	21	18	1	1.948 (0.341)	0.629 (0.166)	0.100 (0.100)	0.369 (0.087)	0.767 (0.195)	70	1		100% Taganrog	216 (Xgpw4082)
	G4	26	16	-	1.567 (0.215)	0.455 (0.138)	0.100 (0.100)	0.266 (0.082)	0.698 (0.234)	60	1		100% Mekki	193 (Xgwm413) 321 (Xgwm372)
	G5	10	8	2	1.487 (0.196)	0.424 (0.126)	0.100 (0.100)	0.244 (0.073)	0.768 (0.194)	70	2		100% Richi	224 (Xgpw4004)
	G6	9	9	3	2.379 (0.274)	0.937 (0.110)	0.078 (0.078)	0.529 (0.051)	0.893 (0.107)	100	7		41% Roussia 59% Souri	-
	G7	22	6	3	1.152 (0.077)	0.166 (0.075)	0.000 (0.000)	0.103 (0.049)	1.000 (0.000)	40	0		100% Badri	232 (Xgpw4082)
	G8	10	9	3	1.905 (0.183)	0.733 (0.103)	0.010 (0.010)	0.428 (0.056)	0.962 (0.038)	100	4		100% Azizi	-
	G9	37	27	3	1.799 (0.219)	0.609 (0.139)	0.043 (0.043)	0.362 (0.080)	0.911 (0.080)	80	2		100% Biskri	-
	Total	302	188	-	1.892 (0.092)	0.620 (0.047)	0.072 (0.022)	0.346 (0.024)	0.835 (0.042)	75,83 (5,43)	-	0.259 (0.079)	-	-
Regions	Gabes	38	31	3	3.031 (0.491)	1.296 (0.122)	0.056 (0.047)	0.610 (0.045)	0.879 (0.11)	100	17			171 (Xgwm193)
	Kairouan	67	25	6	2.707 (0.405)	1.048 (0.136)	0.042 (0.041)	0.563 (0.054)	0.880 (0.117)	100	2			-
	Mahdia	27	21	4	2.883 (0.293)	1.275 (0.102)	0.081 (0.077)	0.619 (0.04)	0.873 (0.121)	100	11			-
	Mednine	22	7	3	1.960 (0.158)	0.790 (0.099)	0.095 (0.095)	0.45 (0.056)	0.851 (0.149)	100	1			-
	Sousse	9	7	1	1.366 (0.185)	0.305 (0.125)	0.100 (0.100)	0.183 (0.073)	0.691 (0.218)	50	1			191 (Xgwm413) 223 (Xgwm285) 224 (Xgpw4004)
	Total	163	91	17	2.389 (0.168)	0.943 (0.073)	0.075 (0.033)	0.485 (0.033)	0.851 (0.059)	90 (10)	-	1.037 (0.239)		-
Climatic stages	High-arid	67	25	6	2.707 (0.405)	1.050 (0.136)	0.042 (0.041)	0.563 (0.054)	0.880 (0.117)	100	2			-
	Low semi-arid	36	28	5	3.006 (0.356)	1.283 (0.113)	0.086 (0.083)	0.622 (0.046)	0.870 (0.126)	100	12			-
	Mid-arid	60	38	6	3.174 (0.433)	1.318 (0.109)	0.070 (0.065)	0.642 (0.039)	0.870 (0.122)	100	19			-
	Total	163	91	17	2.962 (0.225)	1.216 (0.071)	0.066 (0.036)	0.609 (0.027)	0.874 (0.068)	100	-	3.813 (0.571)		-

Acc : Number of accessions ; **MLG :** Number of Multi Locus Genotypes ; **S :** Number of sites ; **Ne :** Number of Effective Alleles ; ***I*** **:** Shannon's Information Index ; ***Ho*** **:** Observed Heterozygosity ; ***He*** **:** Expected Heterozygosity ; ***Fis*** **:** Fixation Index ; **P :** Percentage of Polymorphic Loci ; **PA :** Number of Private Alleles ; ***Nm*** **:** gene flow ; **Var-Pop :** Name of the variety-population ; **DA :** Diagnostic alleles ; **# :** Frequence (0.7-1). ***:** a DA is a rare allele with a frequence >70% for a population or region and <30% for the others

Table 4. Analysis of molecular variance (AMOVA) of Tunisian durum wheat accessions using 10 SSR markers by subpopulations as defined by STRCUTURE, by regions and by climatic stages.

	Source	df	SS	MS	Est. Var.	%
Subopulations*	Among	10	1951.085	195.108	8.430	59
	Within	250	1471.172	5.885	5.885	41
	Total	260	3422.257	-	14.314	100
Regions	Among	4	353.123	88.281	2.605	19
	Within	158	1736.681	10.992	10.992	81
	Total	162	2089.804	-	13.597	100
Climatic stages	Among	2	158.647	79.323	1.276	10
	Within	160	1931.157	12.070	12.070	90
	Total	162	2089.804	-	13.346	100

df: degree of freedom ; **SS :** Sum of Squares ; **MS :** Mean Squares ; **% :** pourcentage of variance

*Admix genetic group was excluded from the analysis

Network analysis

The genetic relatedness between genotypes was tested using the mininum spanning network (MSN) based on Bruvo's distance. MSN separated all the accessions into two main clusters (**Figure 3**). The first cluster named C1 grouped accessions belonging to Azizi G1 and G8, Jneh Khotifa G2, Richi G5, Souri and Rousia G6, Badri G7 and Biskri G9, while the second cluster named C2 grouped accessions belonging to Taganrog G3, Mekki G4, Biada G10 and Mahmoudi G11. In addition, the pairwise Nei's genetic distances calculated between the 11 genetic groups were also in agreement with the accession clustering by the MSN (**Table S6**). The highest Nei's genetic distance value (2.416) was recorded between G10 and G5, followed by the genetic distance value (2.319) recorded between G10 and G7. The lowest genetic distance was registered between G1 and G8 (0.421), between G11 and G3 (0.630), and between G3 and G4 (0.630); indicating that G1 and G8, as well as G11, G3 and G4 were the most genetically related groups respectively. In addition, a morphological comparison between the network groupings revealed a significant difference (*p-values*< 0.05) between C1 and C2 for spike shape, spike length, awn length, grain color, grain form, the number of spikelet/spike and for awns and glumes colors (**Table 5**). The cluster C1 had a higher gene diversity (*He*=0.740) and phenotypic diversity (*H'*=0.77) than cluster C2 with *He*=0.425 and *H'*=0.61 (**Table S7 and S8**). The C1 cluster presented higher spike shape, and spike length values than C2; while C2 had significantly higher awns length and grain size traits (**Table 5**).

Table 5. Means of morphological traits calculated for Azizi and Mahmoudi accessions from the center and the south of Tunisia and for all accessions from C1 and C2 clusters. Means with distinct letters show significant differences at 5% threshold between center and southern accessions.

	Center		South		C1	C2
	AZ	MH	AZ	MH		
SC	1 ^{a}	1 ^{a}	1 ^{a}	1 ^{a}	1 ^{a}	1 ^{a}
SS	7 ^{a}	1 ^{b}	7 ^{a}	1 ^{b}	5-7 ^{a}	3 ^{b}
SD	3-5 ^{a}	7 ^{b}	3-5 ^{a}	5 ^{c}	7 ^{a}	7 ^{a}
SL	5 ^{a}	1-3 ^{b}	5 ^{a}	1-3 ^{b}	3-5 ^{a}	1-3 ^{b}
AL	3 ^{a}	5 ^{b}	3 ^{a}	5 ^{b}	3 ^{a}	5 ^{b}
AC	4 ^{a}	3 ^{a}	4 ^{a}	3 ^{a}	4 ^{a}	3 ^{b}
NS	3 ^{a}	2 ^{a}	2 ^{a}	2 ^{a}	2 ^{a}	2 ^{b}
GlC	1 ^{a}	1 ^{a}	1 ^{a}	1 ^{a}	1 ^{a}	2 ^{b}
GC	5 ^{a}	1 ^{b}	5 ^{a}	1 ^{b}	3 ^{a}	1 ^{b}
GF	2 ^{a}	3 ^{b}	2 ^{a}	3 ^{b}	2 ^{a}	3 ^{b}
GS	5 ^{a}	7 ^{b}	5 ^{a}	7 ^{b}	5 ^{a}	5-7 ^{a}
GN	2 ^{a}	3 ^{b}	2 ^{a}	3 ^{b}	2 ^{a}	2 ^{a}

Center: Mahdia, Sousse and Kairouan; **South:** Gabes and Medenine

AZ: Azizi landrace (G1 and G8)**; MH:** Mahmoudi landrace (G11)

C1: Cluster 1 = G1, G2, G5, G6, G7, G8 and G9**; C2:** Cluster 2 = G3, G4, G10 and G11

SC : Spike Color ; **SS :** Spike Shape ; **SD :** Spike Density ; **SL :** Spike Length ; **AL :** Awns Length ; **AC :** Awns Color ; **NS :** Number of Spikelets/Spike ; **GlC :** Glumes Color ; **GC :** Grains Color ; **GF :** Grains Form ; **GS :** Grains Size ; **GN :** Number of Grains/Spikelet ; **Hd :** Heading (days)

Diversity analysis by regions and climatic stages

Morphological diversity analysis by regions and climatic stages

Shannon-Weaver index (*H'*) was assessed based on 12 spike's morphological traits at the five regions (Sousse, Mahdia, Kairouan, Gabes and Medenine) and the three climatic stages (low semi-arid, high-arid and mid-arid climates) of the designated sites (**Table 1**). Kairouan had the highest diversity index (*H'*=0.74) followed by Medenine (*H'*=0.66), while Sousse had a null diversity index indicating no phenotypic variability between accessions in that region where Richi was the only landrace identified. The most polymorphic characters by regions were spike length (*H'*=0.69), grain form (*H'*=0.65), grain color (*H'*=0.62) and number of spikelets/spike (*H'*=0.61). Furthermore, high-arid climate had the highest diversity index (*H'*=0.74), as this climate stage is represented by Kairouan; while low semi-arid climate represented by Mahdia and Sousse had the lowest diversity index (*H'*=0.59). The most polymorphic characters by climatic stages were awn length (*H'*=0.90), grain form (*H'*=0.82), grain color (*H'*=0.79), and number of spikelets/spike (*H'*=0.73).

Polymorphism level of some characters differed distinctly among regions excluding Sousse where an overall homogeneity of the morphological traits was recorded. Awns color was variable among regions and ranged between 0.12 and 0.73, for Kairouan and Mahdia, respectively. Similarly, the highest spike length was registered in Mahdia (0.99), while the least spike length was recorded in Gabes (0.49). Notwithstanding, the Mahdia region was characterized by the lowest spike color and glumes color indices with 0.00 and 0.48, respectively. However, the highest values of spike color and glumes color were recorded in Mednine (0.53) and Kairoun (0.97), respectively. Morphological traits were also variable from one climate stage to another. The low semi-arid climate is characterized by the lowest records for spike length (0.0) and for glumes color (0.41), in contrast to the high-arid climate where spike length and glumes color were the highest (0.48 and 0.96, respectively). Contrary, awns color registered the least value in the high-arid climate (0.12) and the highest record in the mid-arid climate (0.71). However, no variations were observed between regions for grain color and between climate stages for number of grains/spikelet.

In addition, a dominant phenotypic class of some characters was observed among regions (within more than 70% of accessions), except for Sousse which didn't show any differences in morphological traits. In Gabes, long (>9 cm) (84%) and lightly colored (92%) spikes with cylindrical shape (79%), awns shorter than the spike (84%), moderately long grains form (82%) with a small grain size (<0.3 cm) (82%) and an intermediate number of grains/spikelets (2 to 3) (79%) were noted while spikes with medium length (6 to 9 cm) (73%) were dominant in Medenine. For Mahdia spikes with equal awns and spike length (72%) and with small grain size (78%) were largely observed. However, Kairouan was dominated by spikes with awns longer than the spike (72%). Concerning climatic stages, small grain size (<0.3 cm) (72%) were dominant in mid-arid climate zone, whilst, dark colored (72%) spikes with black awns (96%) were dominant in high-arid climate zone. No particular phenotypic classes were observed within the low semi-arid climate (**Table S9**).

Genetic diversity analysis by regions and climatic stages

The analysis of variance showed that 19% and 10% of the total genetic diversity were observed among regions and among climatic stages, respectively, while 81% and 90% of the genetic variabilities were explained by differences within regions and within climatic stages, respectively (**Table 4**).

Genetic diversity by region showed a number of effective alleles ranging from 1.366 for Sousse to 3.031 for Gabes (**Table 3**). Overall and among all investigated regions, Sousse region has shown the lowest genetic diversity indexes, in contrast to the outstanding indexes registered at Gabes. In fact, Gabes region had the highest number of MLG (31) and the highest Shannon's diversity index with 1.296, while Sousse and Medenine had the lowest number of MLG (7), and the lowest Shannon's diversity index registered at Sousse (0.305). Moreover, the percentage of polymorphism was 100% for all regions except for Sousse which was 50%. Gabes had also the highest number of private alleles (17), while Sousse and Medenine had the lowest number of private alleles (1). The fixation index was above 0.800 in each region except for Sousse which was 0.691. Interestingly, solely the diagnostic allele and heterozygosity index were high in Sousse compared to the other regions. In fact, three diagnostic alleles with a frequency that exceeded 70% and a *Ho* of 0.100 were registered at Sousse, compared to only one at Gabes.

The SSR data analysis by climatic stages revealed that the mid-arid climate was outstanding among the studied climatic stages and had the highest number of effective alleles (3.174), the highest Shannon's diversity index (1.318) and the highest number of private alleles (19). Contrary to the mid-arid climate, the high-arid climatic stage showed the lowest number of effective alleles (2.707), the lowest Shannon's diversity index (1.050) and the lowest number of private alleles (2). However, the fixation index was similar among all studied climatic stages and recorded an index above 0.800 for all climatic regions (**Table 3**).

Correlations between genetic distance and geographic distance

The Mantel test showed a significant correlation at (P=0.010; R_{xy}=0.286) between genetic and geographic distances among durum wheat accessions, suggesting that geographically close individuals were genetically related. With the exception of the Sousse region, Azizi and Mahmoudi were the most widespread landraces between central and southern Tunisia across all regions and climatic stages. However, Azizi was more frequent in Gabes (25 accessions out of 38), while Mahmoudi was mostly found in Medenine (13 accessions out of 22) and Mahdia (11 accessions out of 27) (**Figure 4**). In addition, Sousse grouped all G5 genotypes, corresponding to Richi landrace. Kairouan grouped all G7 and G2 genotypes, corresponding to landraces Badri and Jneh Khotifa, respectively. The landrace Taganrog, representative of the genetic group G3, was exclusively found in Mahdia.

Furthermore, and exclusively for the widespread landraces Azizi and Mahmoudi, a comparison of morphological traits between Azizi and Mahmoudi accessions collected from central and southern Tunisia was carried out and has revealed a non-significant difference (*p-values* > 0.05) for all traits, except for spike density within Mahmoudi (*p*=0.00). In fact, Mahmoudi accessions from central Tunisia had compact (SD=7) spikes compared to southern Mahmoudi accessions characterized by lax spikes (SD=5) (**Table 5**).

Discussion

Genetic and morphological diversity kept in Tunisian durum wheat germplasm

In the present study, we investigated the genetic diversity of 302 Tunisian durum wheat accessions using ten SSR markers which were able to reach the maximal differentiation among the multi-locus genotypes suggesting that these markers have a good resolution power (data not shown). Overall, the studied collection is characterized by a high genetic diversity level with an overall number of alleles per locus Na of 9.9, a Polymorphic Information Content PIC of 0.690 and a gene diversity *He* of 0.346. Similar levels of polymorphism (Na=8; PIC=0.68) was previously reported on a Tunisian durum wheat collection composed by 7 modern cultivars and 27 old cultivars fingerprinted with 15 SSR markers [7]. More recently, Slim *et al.* [19] analyzed the genetic diversity of Tunisian durum wheat germplasm composed of 41 traditional varieties and 13 cultivars using 16 SSR markers, showing a mean PIC value of 0.57 and a *He* varying from 0.28 to 0.82, with a number of alleles ranging from 2 to 13. A higher level of polymorphism (Na=10; *He*=0.71) was reported in a wider geographical collection of 172 durum wheat landraces collected from 21 Mediterranean countries and 20 modern cultivars genotyped by 44 SSR markers [22]. However, lower genetic diversity was observed in 33 Anatolian, 136 south Italian and 40 North-West African durum wheat landraces using 14, 44 and 29 SSR markers, respectively [2, 12, 23]. Low genetic diversity (PIC=0.1; *He*=0.25) was also observed in 196 Tunisian durum wheat accessions by Robbana *et al.* [18], due i) to the use of bi-allelic DArTseq markers with lower informativeness level than multi-allelic SSR markers and ii) to a germplasm collection limited to 5 landraces. This variability between results suggests that capturing the maximum genetic diversity would depend essentially on the type of deployed markers, the number of landraces, the origin and geographical distribution (genetic backgrounds) of the studied collection.

Based on 12 morphological traits, the levels of phenotypic diversity detected in our study were consistent with those observed for genetic diversity, with a Shannon-Weaver index *H'* of 0.80. The morphological diversity was higher than the previously described for Tunisian durum wheat germplasm (*H'*=0.53) of 930 accessions collected from a reduced number of sites from southern Tunisia, using twenty-two qualitative and three quantitative traits [16]. Lower phenotypic diversity was also observed for Moroccan durum wheat populations composed of 101 landraces (*H'*=0.62) using nine agro-morphological traits [24] and of 59 traditional durum wheat (*H'*=0.78) using nine agro-morphological traits [25]. Ethiopian durum wheat populations composed of 32 landraces had an *H* index of 0.71 using 8 qualitative traits [26], while Oman populations composed of 161 accessions showed *H'* index of 0.52 and 0.66 using 15 qualitative and 17 quantitative characters respectively [26].

In this study, spike length (*H'*=0.98), grain size (*H'*=0.94), grain form (*H'*=0.87), grain color (*H'*=0.86) and spike shape (*H'*=0.86) were the most polymorphic morphological traits. Previous studies of Tunisian durum wheat populations showed different results for polymorphic traits based on UPOV and IPGRI. Belhadj *et al.* [16] and concluded that the most polymorphic traits were width of the truncation (*H'*=0.97) and spike color (*H'*= 0.92); whereas Ayed *et al.* [17] revealed that number of grains/spike (*H'*= 0.91), yield (*H'*= 0.89), plant height (*H'*=0.87) and thousand kernel weight (*H'*=0.86) had the highest diversity index. Slim *et al.* [27] reported that polymorphism was high for awn anthocyanin coloration (*H'*=1.18), spike glaucosity (*H'*=0.89), hairiness on the external surface (*H'*=0.88), awn color (*H'*=0.78) and awn length in relation to the spike (*H'*=0.77). Ayed and Slim [28] revealed that spike density (*H'*=0.86), glume pubescence (*H'*=0.80) and glume color (*H'*=0.79) showed the highest diversity index. These differences were essentially related to landraces. Indeed, Ayed *et al.* [17] assessed 17 Tunisian durum wheat landraces, which may have contributed to the wider range of morphological variation, compared to the present study and other studies with a less number of landraces. Thus, increasing the number of landraces would allow for capturing a greater agro-morphological diversity.

Population structure, network analysis and relationships between genetic and morphological data

In addition to farmer's selection pressure for specific types of landraces, natural selection was observed morphologically within a single landrace, Mahmoudi. Mahmoudi accessions collected from southern Tunisia showed significantly looser spikes than Mahmoudi accessions collected from central Tunisia, characterized by compact spikes. We might suggest that the relaxed spike characterized by an open glume in the Mahmoudi accessions originated from the south could provide tolerance to high temperature by maintaining fertility as it has been observed in rice germplasms [36]. This differentiation between Mahmoudi accessions offers potential tools i) to use relaxed spike trait in breeding programs for heat stress tolerance, and ii) to identify genes and mechanisms involved in flower development useful for improving wheat adaptation to arid and marginal environments.

MSN analysis grouped the accessions into two major clusters C1 and C2. However, C1 and C2 do not correlate with the landraces' geographical origins. Notably Mahmoudi and Biskri were both introduced from Algeria, while landraces Jneh Khotifa, Azizi, Mekki, Biada and Roussia were considered as local landraces that were cultivated mainly in the north and the center of Tunisia. Nevertheless, landraces Azizi and Mekki had various reported origins, however, no origin has been attributed to landraces Richi and Taganrog that were reported as very old landraces but not local [9, 10]. According to Deghais *et al.* [10], the landrace Jneh Khotifa was also called Jneh Zarzoura and/or Kahla; the denomination of the landrace Souri was extended, from 1915, to Sarebouza wheat received from Armenia. Soriano *et al.* [22], using 44 SSR, reported that Tunisian durum wheat landraces have four geographical origins, namely East Mediterranean, East Balkan and Turkey, West Balkan and Egypt, and West Mediterranean, with dominance (at more than 50%) of the West Mediterranean genetic group. In addition, Sorriano *et al.* [22] demonstrated that western Mediterranean landraces were characterized by the heaviest grain weight compared to the three other genetic groups. In our study, grain size did not significantly differ between C1 and C2 suggesting that both clusters have indeed western Mediterranean origin. Moreover, Robbana *et al.* [18] mentioned that most of Tunisian landraces were introduced from the early Carthage trade maritime activity in the Mediterranean Sea, through pathways from Lebanon, Greece and Italy.

Conclusions

Tunisian old durum wheat, characterized here by both high genetic and morphological diversity, represents an important and valuable genetic resource that should be included in breeding and well-established conservation programs. This study showed that Tunisian old durum wheat is structured into landraces revealing the effect of selection pressure directed by farmers for specific wheat types and agro-morphologies. Nevertheless morpho-geographical spike density trait revealed specifically in Mahmoudi accessions suggests that environmental selection may occur. Thus, our results provide an interesting venue to improve wheat adaptation to extreme or fluctuating Mediterranean conditions. Further physiological and agronomic analysis should be conducted to ascertain whether this trait could be exploited in durum wheat breeding programs for tolerance to heat and drought.

Methods

Local durum wheat collection and multiplication

A collection of 304 old durum wheat accessions provided by the National Gene Bank of Tunisia (BNG) was used for this study. Accessions were collected from five regions that belong to three distinct climatic zones - central Tunisia, which is characterized by a low semi-arid climate at Sousse and Mahdia regions and by a high-arid climate at Kairouan region, and southern Tunisia, which is characterized by a mid-arid climate at Gabes and Mednine regions. Global Positioning System (GPS) coordinates of 163 out of the 304 accessions were registered (**Table S10**). Each accession was sown and purified from a single spike-derived lineage by the BNG team and a BNG code has been assigned to each accession. All accessions were further multiplied for spike characterization. The collected set of accessions, used in this study, is preserved at the BNG of Tunisia and is available upon request.

DNA extraction and SSR genotyping

Five seeds from one spike of each accession were germinated and grown under controlled conditions with a photoperiod of 16h/24h, a hygrometry (RH) of 70% and a temperature of 20°C/16 °C (day/night rhythm) at Bioger research unit, INRAE France. At the seedling growth stage (Zadock scale 13-14), one leaf of each accession was sampled and placed in extraction plates. The plates were placed at (-80°C) for 12h before DNA extraction. DNA extraction for each of the 304 accessions was carried out using the DNeasy PowerPlant Pro HTP 96 Kit (Qiagen, France). DNA concentrations were quantified using a Nanodrop spectrophotometer (ND-1000) and stored at (-20°C) for subsequent processing. For each accession, DNA was adjusted to 15 ng.µl^{-1} and genotyped using 10 SSR markers (**Table S4**). The 10 SSR markers used in this study were selected among 15 SSR markers used in Sahri et al. (2014) [24] study. The forward primers were labeled with fluorescent dyes and SSR markers were multiplexed as described by Gautier *et al.* [37]. For each multiplex, PCRs and electrophoresis were accomplished according to a protocol established by Eurofin (https://www.eurofins.fr). DNA amplification was performed by preheating the DNA at 95°C for 5 mn, followed by 35 cycles of 95°C for 30 s, 60°C for 90 s and 72°C for 30 s, with a final extension step of 60°C for 30 mn. PCR products were checked for amplification on a 2% agarose gel and fragments were separated according to their size on an ABI Prism Genetic Analyzer (Applied Biosystems). Data was checked again using Peak scanner software version 1.0. Two accessions had missing data for all used SSR and were discarded from our study.

Morphological characterization

The morphological characterization was carried out on five spikes per accession. Overall, a total of 1520 spikes were characterized among the entire studied collection. Accessions were evaluated using 12 quantitative and qualitative spike morphological traits. Spike evaluation was based on durum wheat descriptor standards of the International Plant Genetic Resources Institute [20] and the International Union for the Protection of New Varieties of Plants [21] (**Table S11**). Spike morphological traits, defined by distinct phenotypic classes, were visually and numerically estimated. Visual phenotypic estimates were attributed to the spike (SC), glumes (GIC), awns (AC) and grains (GC) colors, the density (SD) and the shape (SS) of the spike, the grains form (GF) and size (GS), and awns length compared to the spike. However, the grain size (GS), spike length (SL), the number of spikelets per spike (NS) and the number of grains per spikelet (GN) were measured for each accession and then converted into codes.

Subsequently, accessions were named based on the catalog of cereal varieties cultivated in Tunisia [10]. In fact, the catalog represents a reference reporting and describing typical varietal characteristics of more than 40 old local durum wheat landraces cultivated in Tunisia. The first author of this paper, PhD student, has undertook the formal identification and morphological characterization of the 304 accessions used in the present study

Data analysis

Polymorphism of SSR markers using polymorphic information content PIC

To measure the informativeness of the markers, the average polymorphic information content (PIC) was calculated for each SSR by determining the frequency of alleles per locus according to the formula given by Powell *et al.* [38]:

$$\mathbf{PIC} = 1 - \sum_{i=1}^{n} f_i^2$$

Polymorphism of morphological traits using Shannon-Weaver index H'

Frequencies of the different phenotypic classes were calculated for each of the 12 spike's morphological traits in the whole collection, by regions, by climatic stages (**Table S9**) and by landraces (**Table S1**). Based on these frequencies, the Shannon-Weaver index (*H*) was calculated for each trait using Past software [39]. *H* was estimated for the entire durum wheat collection, regions, climatic stages and for each landrace. Each value of *H* was standardized by conversion to a relative phenotypic diversity index (*H'*) in order to express the values of *H'* in the range of 0-1. Index *H'* was calculated as follows:

$H' = H / H_{max}$

Where H_{max} = Ln (S), S = number of phenotypic classes

Morphological relationship between accessions and population genetic structure analysis

A principal coordinate analysis (PCoA) was performed based on the 12 spike's morphological traits to investigate the empirical phenotypic distances between the 304 accessions using R 3.3.2 [46].

Based on SSR data generated for 302 accessions, 188 multilocus genotypes (MLG) were identified with GIMLET software version 1.3.2 [41] .A population genetic structure analysis was conducted on the 188 MLG, using the program STRUCTURE version 2.3.4 [42]. The run was conducted with K-values varying from 1 to 20 in an admixture ancestry model applying 10 independent runs for each of the different K values. A burn-in phase of 100,000 iterations and 100,000 Markov Chain Monte Carlo (MCMC) iterations were performed. The run with maximum likelihood was used to assign individual genotypes into genetic groups. Genotypes with affiliation probabilities (inferred ancestry) > 75% were assigned to a distinct genetic group and those with < 75% were treated as admixed. Plot of mean posterior probability (ln P(D)) values per clusters (K), and delta-K method of ln P(D) STRUCTURE harvester version 0.6.94 were used to determine the optimal number of genetic groups [43] .

In addition, a minimum spanning network (MSN) based on Bruvo's distance [44] (Bruvo *et al.,* 2004) using "*poppr*" and "*adegenet*" packages was generated under R 3.3.2 [45], in order to classify the 302 accessions according to their genetic relationship . Furthermore, a mean of each of the 12 spike's morphological traits was calculated for accessions belonging to the different clusters defined by the MSN analysis, as follows:

$$\mathbf{Mean} = \sum_{i=1}(nC_i)/N$$

Where N is the number of genotypes per genetic cluster as defined by the MSN analysis, n is the number of individuals per phenotypic class and C_i is the i^{th} phenotypic class per morphological trait.

Based on the calculated means, an analysis of variance ANOVA was carried out using R 3.3.2 [46] to test for significant differences between genetic clusters for each morphological trait.

Population genetics and data analysis by regions and climatic stages

GenAlEx version 6.501 [40] was used to calculate the number of alleles (Na), the number of effective alleles (Ne), the number of private alleles (PA : alleles specific to a single population), the Shannon's information index (*I*), the expected (*He*) and observed (*Ho*) heterozygosity, the fixation index (*F*), the percentage of polymorphic loci (P), and the diagnostic alleles (DA is a rare allele with a frequence >70% for a genetic group or region and <30% for the others) within each genetic group, region and climatic stage.

In addition, the correlation between the genetic distance and the log (1+geographic distance) transformed geographic distance of accessions was analyzed using a Mantel test [46] for the entire collection, under GenAlEx version 6.501. Correlations between the genetic distance matrix and morphological distance matrix were also assessed using a Mantel test.

Furthermore, an analysis of molecular variance (AMOVA) was performed under GenAlEx version 6.501 to investigate the significance of genetic group differentiation as defined by STRUCTURE and the genetic variability explained by regions and climatic stages.

Moreover, a mean of the 12 spike's morphological traits was estimated for Azizi and Mahmoudi accessions, both existing in different climatic zones of central and southern Tunisia, as described above. For each morphological trait, an analysis of variance ANOVA was carried out using R 3.3.2 [45] to test for potential regional effects on the morphological traits.

Abbreviations

AMOVA: Analysis of molecular variance; C: Genetic cluster as defined by MSN analysis; G: Genetic group as defined by STRUCTURE; *H'*: Shannon-Weaver index; MSN: Minimum spanning network; PCoA: Principal component analysis; PIC: Polymorphic information content.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The data sets supporting the results of this article are included in this manuscript and its additional information files.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the federated project entitled “Identification of durum wheat resistant genotypes to biotic and drought stress and their valorization for sustainable agriculture” acronym RESIDUR, supported by IRESA under the Tunisian Ministry of Agriculture.

Authors Contributions

Conceptualization and Supervision of the study were performed by SH. MO, SF and TM realized the experiments and participated in genotyping. MM assembled the panel. BB and MO carried out the data analysis. MO, BB and SH participated in interpreting the data. MO prepared and wrote the original draft. Revising and editing the manuscript were performed by BB and LA. All co-authors approved the final version of the manuscript

Acknowledgments

Not applicable

References

1. Feldman M. Origin of cultivated wheat. In: Bonjean AP, Angus WJ, editors. The world wheat book: A history of wheat Breeding. Lavoisier, Paris. 2001. p. 3–57.
2. Royo C, Nazco R, Villegas D. The climate of the zone of origin of Mediterranean durum wheat (*Triticum durum Desf.*) landraces affects their agronomic performance. Genet Resour Crop Ev. 2014; 61: 1345–
3. Royo C, Maccaferri M, Álvaro F, Moragues M, Sanguineti MC, Tuberosa R, Maalouf F, García del Moral LF, Demontis A, Rhouma S, Nachit M, Nserallah N, Villegas D. Understanding the relationships between genetic and phenotypic structures of a collection of elite durum wheat accessions. Field Crops Res. 2010; 119: 91–
4. Alsaleh A, Baloch FS, Nachit M, Özkan H. Phenotypic and genotypic intra-diversity among Anatolian durum wheat “Kunduru” landraces. Syst. Ecol. 2016; 65: 9–16.
5. Bœuf F. Le blé en Tunisie. Annales du service botanique et agronomique de Tunisie, Tunis, Tunisia. 1932;18–43.
6. Ayed S, Amara HS. Distribution and phenotypic variability aspects of some quantitative traits among durum wheat accessions. Crop Sci. J. 2009; 16: 219–224.
7. Medini M, Hamza S, Rebai A, Baum M. Analysis of genetic diversity in Tunisian durum wheat cultivars and related wild species by SSR and AFLP markers. Genet Resour Crop Ev. 2005; 52: 21–31.
8. Gharbi MS, Deghais M, Ben Amar F. Breeding for resistance to *Septoria tritici* in durum wheat. In: Royo C, Nachit M, Di Fonzo N, Araus JL, editors. Durum wheat improvement in the Mediterranean region: New challenges. Zaragoza, CIHEAM. 2000. p. 397–401.
9. Ammar K, Gharbi M.S, Deghais M. Wheat in Tunisia. In: Bonjean AP, Angus WJ (Eds) The world wheat book: A history of wheat Breeding. Lavoisier, Paris. 2011;443–465.
10. Deghais M, Kouki M, Gharbi MS, El Felah M. Les Variétés de Céréales Cultivées en Tunisie (blé dur, blé tendre, orge et triticale). INRAT Eds, Tunis, Tunisia. 2007.
11. Ben Salem M, Daaloul A, Ayadi A. Le blé dur en Tunisie. In: Durum wheat quality in the Mediterranean region. Zaragoza, CIHEAM. 1995; 81–91.
12. \Marzario S, Logozzo G, David JL, Spagnoletti Zeuli P, Gioia T. Molecular Genotyping (SSR) and Agronomic Phenotyping for Utilization of Durum Wheat (*Triticum durum Desf.*) Ex Situ Collection from Southern Italy: A Combined Approach Including Pedigreed Varieties. Genes. 2018; 9: 465.
13. Lopes MS, El-Basyoni I, Baenziger PS, Singh S, Royo C, Ozbek K, Ban T. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. Exp. Bot. 2015; 66: 3477–3486.
14. Soriano JM, Villegas D, Sorrells ME, Royo C. Durum Wheat Landraces from East and West Regions of the Mediterranean Basin Are Genetically Distinct for Yield Components and Phenology. Plant Sci. 2018; 9: 80.
15. Baraket G, Abdallah D, Mustapha SB, Tamarzizt HB, Salhi-Hannachi A. Combination of simple sequence repeat, S-Locus polymorphism and phenotypic data for identification of Tunisian plum species (*Prunus*). Biochem Genet. 2019; 57: 673-694.\
16. Belhadj H, Medini M, Bouhaouel I, Amara HS. Analyse de la diversité phénotypique de quelques accessions autochtones de blé dur (*Triticum turgidum ssp. durum Desf.*) du sud tunisien. New. Sci. Agri. Biotech. 2015; 24 :1115-1125.
17. Ayed S, Karmous C, Trifa Y, Slama AO, Amara HS. Phenotypic diversity of Tunisian durum wheat landraces. Crop Sci. J. 2010; 18: 35–42.
18. Robbana C, Kehel Z, Naceur B, Sansaloni C, Bassi F, Amri A. Genome-Wide Genetic Diversity and Population Structure of Tunisian Durum Wheat Landraces Based on DArTseq Technology. J. Mol. Sci. 2019; 20: 1352.
19. Slim A, Piarulli L, Chennaoui KH, Rouaissi M, Robbana C, Chaabane R, Mangini G. Genetic Structure Analysis of a Collection of Tunisian Durum Wheat Germplasm. J. Mol. Sci. 2019; 20: 3362.
20. Descriptors for wheat. International Board for Plant Genetic Resources. Rome, Italy. 1985.
21. Union Internationale pour la Protection des Obtentions Végétales (UPOV). Principes directeurs pour la conduite de l’examen des caractères distinctifs, de l’homogénéité et de la stabilité, blé dur. 1988 ; 3–32.
22. Soriano JM, Villegas D, Aranzana MJ, García del Moral LF, Royo C. Genetic Structure of Modern Durum Wheat Cultivars and Mediterranean Landraces Matches with Their Agronomic Performance. PLoS One. 2016;
23. Oliveira HR, Campana MG, Jones H, Hunt HV, Leigh F. Tetraploid wheat landraces in the mediterranean Basin: taxonomy, Evolution and genetic diversity. PLoS One. 2012 ; 7: 1–13.

24. Sahri A, Chentoufi L, Arbaoui M, Ardisson M, Roumet P, Belqadi L, Birouk A, Roumet P, Muller MH. Towards a comprehensive characterization of durum wheat landraces in Moroccan traditional agrosystems: analysing genetic diversity in the light of geography, farmers' taxonomy and tetraploid wheat domestication history. BMC Evol Biol. 2014 ; 14 : 264.
25. Chentoufi L, Sahri A, Arbaoui M, Belqadi L, Birouk A, Roumet P, Muller MH. Diversité agro-morphologique et gestion variétale par les agriculteurs du blé dur (*Triticum turgidum ssp. durum*) dans le Pré-Rif marocain. rev mar sci agron vét. 2014 ; 2: 30–38.
26. Al Khanjari S, Filatenko AA, Hammer K, Buerkert A. Morphological spike diversity of Omani wheat. Genet Resour Crop Ev. 2008; 55: 1185–1195.
27. Slim A, Ayed S, Slama AO, Robbana C, Jaime AT, Slim-Amara H. Morphological Diversity of Some Qualitative Traits in Tetraploid Wheat Landrace Populations Collected in The South of Tunisia. J. Plant Breed. 2011; 5: 67–70.
28. Ayed S, Slim AH. Distribution and phenotypic variability aspects of some quantitiative traits among durum wheat accessiions. Crop Sci. J. 2008; 16: 4.
29. Wang LX, Jun QIU, Chang LF, Liu LH, LI HB, Pang BS, Zhao CP. Assessment of wheat variety distinctness using SSR markers. Integr. Agric. 2015; 14: 1923–1935.
30. Mengistu DK, Kiros AY, Pè ME. Phenotypic diversity in Ethiopian durum wheat (Triticum turgidum var. durum) landraces. Crop J. 2015 ; 3:190–9.https://doi.org/10.1016/j.cj.2015.04.003.
31. Bezançon G, Pham JL, Deu M, Vigouroux Y, Sagnard F, Mariac C, Kapran I, Mamadou A, Gérard B, Ndjeunga J, Chantereau J. Changes in the diversity and geographic distribution of cultivated millet (*Pennisetum glaucum (L.) R. Br.*) and sorghum (*Sorghum bicolor (L.) Moench*) varieties in Niger between 1976 and 2003. Genet Resour Crop Ev. 2009; 56: 223-236.
32. Jaradat AA. Wheat Landraces: A mini review. Emir J. Food Agr. 2013; 25:20–
33. Abdel-Ghani AH. Genetic diversity and Population Structure of Jordanian Durum Wheat (*Triticum turgidum L. subsp durum*) Landraces as Revealed by RAPD Markers. JJAS. 2013; 9: 369–382
34. Baloch FB, Alsaleh A, Saenz de Miera LE, Hatipoglu R, Çiftçi V, Karakoy T, Yıldız M, Ozkan H. DNA based iPBS-retrotransposon markers for investigating the population structure of pea (*Pisum sativum*) germplasm from Turkey. Syst. Ecol. 2015; 61: 244-252.
35. Fayaz F, Sarbarzeh MA, Talebi R, Azadi A. Genetic Diversity and Molecular Characterization of Iranian Durum Wheat Landraces (*Triticum turgidum durum (Desf.) Husn.*) Using DArT Markers. Biochem Genet. 2018; 57: 98–116.
36. Yan H, Zhang B, Zhang Y, Chen X, Xiong H, Matsui T, Tian X. High temperature induced glume closure resulted in lower fertility in hybrid rice seed production. Front. Plant Sci. 2017; 7: 1960
37. Gautier A, Marcel TC, Confais J, Crane C, Kema G, Suffert F, Walker AS. Development of a rapid multiplex SSR genotyping method to study populations of the fungal plant pathogen *Zymoseptoria tritici*. BMC Res. Notes. 2014: 7, 373.
38. Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germoplasm analysis. Mol Breed. 1996; 2: 225–
39. Hammer Ø, Harper DAT., Ryan PD. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia electronica*. 2001; 4: 9. http://palaeoelectronica.org/2001_1/past/issue1_01.htm
40. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Ecol. Notes. 2012; 6: 288–295. http://dx.doi.org/10.1093/bioin-formatics/bts460.
41. Valiere N. Gimlet: a computer program for analyzing genetic individual identification data. Ecol. Notes. 2002; 2: 377–379.
42. Pritchard JK, Stephens P, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155: 945–959.
43. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: A Website and program for visualizing STRUCTURE output and implementing the Evanno method. Genet. Resour. 2012; 4: 359–361. http://taylor0.biology.ucla.edu/structureHarvester/
44. Bruvo R, Michiels NK, D'Souza TG, Schulenburg H. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Ecol. 2004; 13: 2101-2106.
45. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013 http://www.R-project.org/
46. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res. 1967; 27, 209–220.

Figures

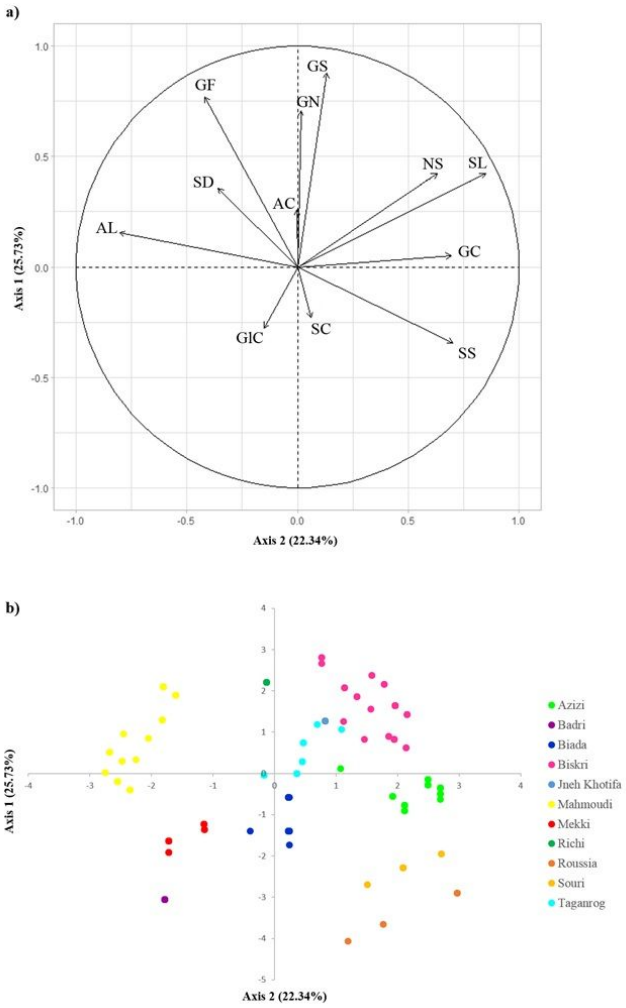


Figure 1

Principal component analysis plot depicting 11 durum wheat landraces within 304 Tunisian accessions using 12 morphological traits under GenAlEx software [40] ; (a) Projection of the 12 variables on the PCoA plot axes. SS : Spike Shape ; SL : Spike Length ; AL : Awns Length ; SC : Spike Color ; NS : Number of Spikelets/Spike ; GlC : Glumes Color ; GN : Number of Grains/Spikelet ; GF : Grains Form ; GS : Grains Size ; GC : Grains Color ; AC : Awns Color ; SD : Spike Density; (b) Projection of the 304 accessions on the PCoA plot axes. Accessions were color-coded according to their landraces nomenclature, as identified with the morphological characterization.

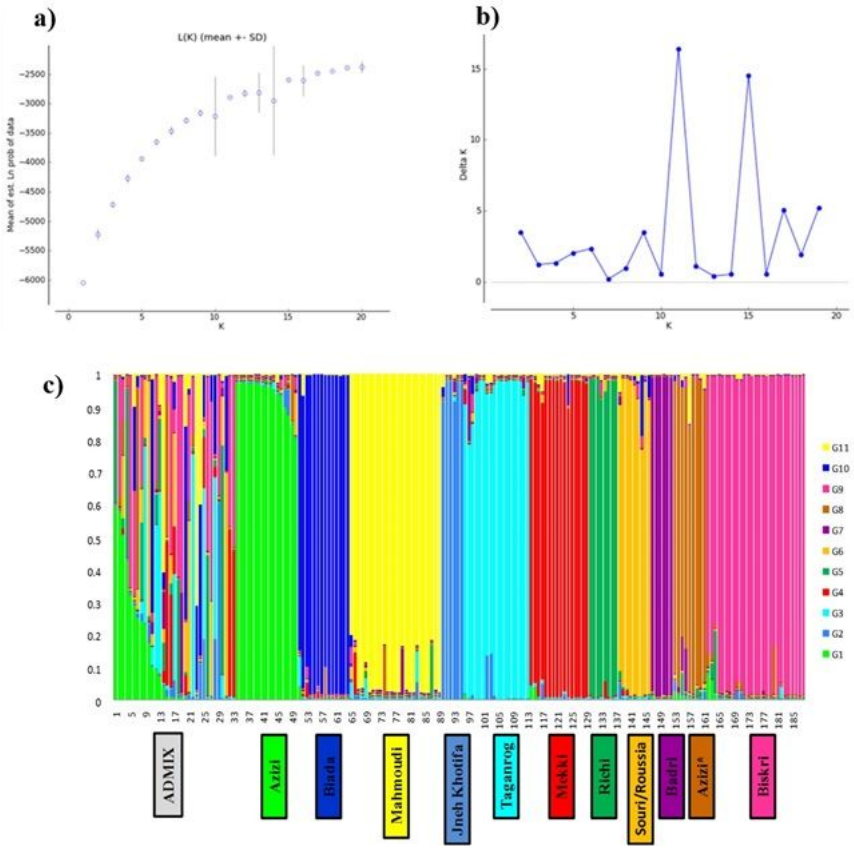


Figure 2

Population structure analysis of 302 Tunisian durum wheat accessions genotyped with 10 SSR markers: (a) Plot of mean posterior probability (ln P(D)) values per cluster (K); (b) delta-K analysis of Ln P(D). STRUCTURE program where used based on 10 replicates per K, for K ranging from 1 to 20, with a burn-in period of 100,000 and Monte Carlo Markov Chain replicates of 100,000 iterations; (c) Membership coefficient bar plot displaying population structure at K = 11 for 302 Tunisian durum wheat accessions genotyped with 10 SSR markers inferred from STRUCTURE [42]. Each MLG is represented by a vertical line and they are ordered by color-coded genetic subpopulation (G1 to G11). For each genetic subpopulation, corresponding durum wheat landrace is mentioned. * Azizi landrace was divided into two genetic subpopulations G1 and G8.

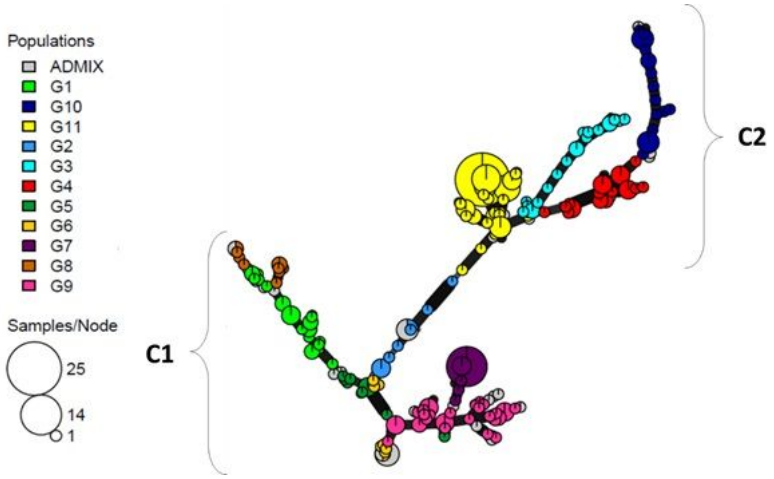


Figure 3

Minimum spanning network using Bruvo's distance of 302 durum wheat accessions genotyped with 10 SSR markers, performed under R software. Each node represents a multilocus genotype (MLG) and the size of the node is proportional to the number of accessions representing the MLG. MLGs were color-coded according to their membership to a genetic subpopulation (G1 to G11) as defined by STRUCTURE at K=11. Admixed individuals were color-coded in gray. Edge widths represent relatedness.

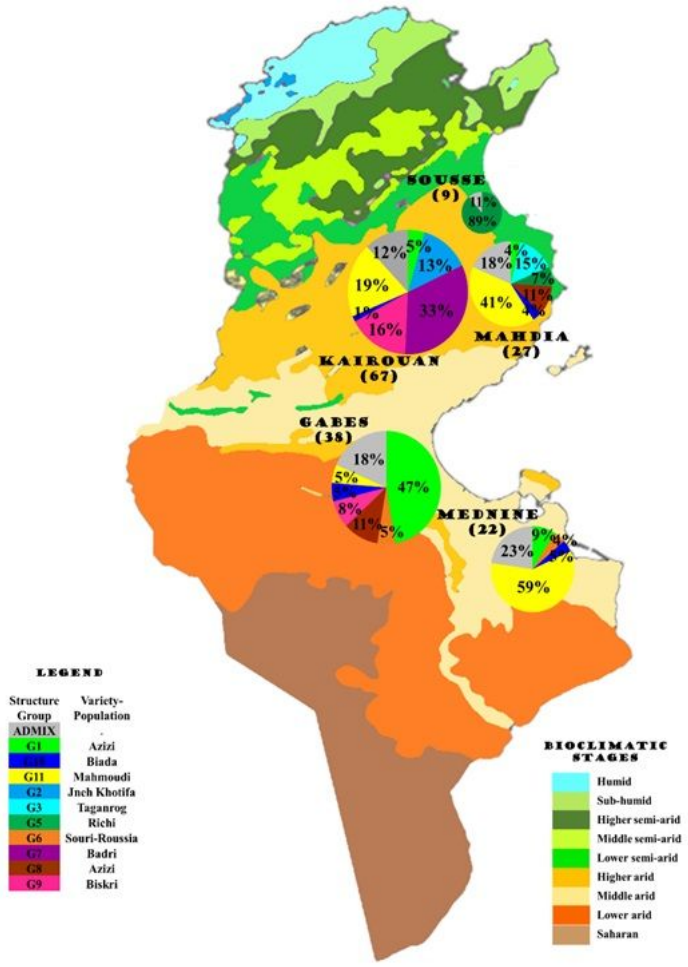


Figure 4

Geographic distribution of the 11 genetic subpopulations (G1-G11), defined by STRUCTURE (Earl 2012) on 163 geo-localized durum wheat accessions genotyped with 10 SSR markers, over the regions of origin and the bioclimatic stages in Tunisia (https://www.d-maps.com/).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- RevisedSupplementaryInformationBMCGenetics.pdf