



Article

Elucidating the Genetic Relationships on the Original Old Sicilian *Triticum* Spp. Collection by SNP Genotyping

Maria Carola Fiore ¹, Sebastiano Blangiforti ², Giovanni Preiti ³, Alfio Spina ⁴, Sara Bosi ⁵, Iliara Marotti ⁵, Antonio Mauceri ³, Guglielmo Puccio ⁶, Francesco Sunseri ³ and Francesco Mercati ^{6,*}

- ¹ Council for Agricultural Research and Economics (CREA), Research Centre for Plant Protection and Certification, 90011 Bagheria, Italy
² Stazione Consorziale Sperimentale di Granicoltura per la Sicilia, Santo Pietro, 95041 Caltagirone, Italy
³ Department AGRARIA, University Mediterranea of Reggio Calabria, 89122 Reggio Calabria, Italy
⁴ Council for Agricultural Research and Economics (CREA), Research Centre for Cereal and Industrial Crops, 190, 95024 Acireale, Italy
⁵ Department of Agricultural and Food Science, Alma Mater Studiorum, University of Bologna, Viale Fanin, 40127 Bologna, Italy
⁶ National Research Council (CNR) of Italy, Institute of Biosciences and Bioresources (IBBR), 90129 Palermo, Italy
* Correspondence: francesco.mercati@ibbr.cnr.it; Tel.: +39-0916574578



Citation: Fiore, M.C.; Blangiforti, S.; Preiti, G.; Spina, A.; Bosi, S.; Marotti, I.; Mauceri, A.; Puccio, G.; Sunseri, F.; Mercati, F. Elucidating the Genetic Relationships on the Original Old Sicilian *Triticum* Spp. Collection by SNP Genotyping. *Int. J. Mol. Sci.* **2022**, *23*, 13378. <https://doi.org/10.3390/ijms232113378>

Academic Editor: Hongjie Li

Received: 25 September 2022

Accepted: 28 October 2022

Published: 2 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Several *Triticum* species spread in cultivation in Sicily and neighboring regions over the centuries, which led to the establishment of a large genetic diversity. Many ancient varieties were widely cultivated until the beginning of the last century before being replaced by modern varieties. Recently, they have been reintroduced in cultivation in Sicily. Here, the genetic diversity of 115 and 11 accessions from Sicily and Calabria, respectively, belonging to *Triticum* species was evaluated using a high-density SNP array. Einkorn, emmer, and spelta wheat genotypes were used as outgroups for species and subspecies; five modern varieties of durum and bread wheat were used as references. A principal coordinates analysis (PCoA) and an unweighted pair group method with arithmetic mean (UPGMA) showed four distinct groups among *Triticum* species and *T. turgidum* subspecies. The population structure analysis distinguished five gene pools, among which three appeared private to the *T. aestivum*, *T. turgidum* subsp. *Turgidum*, and 'Timilia' group. The principal component analysis (PCA) displayed a bio-morphological trait relationship of a subset (110) of ancient wheat varieties and their wide variability within the *T. turgidum* subsp. *durum* subgroups. A discriminant analysis of principal components (DAPC) and phylogenetic analyses applied to the four durum wheat subgroups revealed that the improved varieties harbored a different gene pool compared to the most ancient varieties. The 'Russello' and 'Russello Ibleo' groups were distinguished; both displayed higher genetic variability compared to the 'Timilia' group accessions. This research represents a comprehensive approach to fingerprinting the old wheat Sicilian germplasm, which is useful in avoiding commercial fraud and sustaining the cultivation of landraces and ancient varieties.

Keywords: biodiversity; bio-morphological traits; single-nucleotide polymorphism (SNP); *Triticum*; wheat landraces

1. Introduction

Wheat is one of the key crops to ensure that food and nutritional needs are provided for an ever-growing human population in the future. This is attributable to its ability to adapt to a wide range of climates worldwide. Wheat and its derivatives are foreseen to provide about 20% per capita consumption of protein and calories for an estimated global population of 9.7 billion people in 2050 [1] by contributing to essential components to the human diet [2].

Wheat represents the second most important cereal grain after maize and is the most widely grown crop in the world, with a cultivated area of approximately 219 million hectares and a production of about 760 million tons [3]. Bread or common wheat (*Triticum aestivum* L.) comprises roughly 90% to 95% of the total wheat production worldwide, whereas durum wheat (*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.) represents just 5–7%. However, durum wheat is a staple food in Mediterranean countries due to its adaptation to hot dry conditions and is useful for producing pasta, couscous, bulgur, and bread [4]. Italy is the country leader in EU production and is the second-largest world producer after Canada with a production of 3,885,216 tons [5], a leading role attributable to the economic importance of the pasta industry.

Wheat belongs to the grass family *Poaceae* and the tribe *Triticeae*. Einkorn wheat (*T. monococcum* ssp. *monococcum*, $2n = 2x = 14$, AA) is the earliest diploid species of cultivated wheat. Tetraploid wheat ($2n = 4x = 28$, AABB) includes emmer (*T. turgidum* ssp. *dicoccum*), which was widely cultivated before bread wheat [6], and durum wheat (*T. turgidum* ssp. *durum*), which was cultivated mainly in the Mediterranean basin [7]. Bread wheat (*T. aestivum* L.) is an allohexaploid ($2n = 6x = 42$, AABBDD) that arose from the interspecific hybridization of *T. turgidum* ssp. *dicoccum* (AABB) and *Aegilops tauschii* Coss. (DD) [8].

The most significant effect of domestication first and breeding later has been a reduction in the genetic diversity of cultivated wheat species [9,10]. In Italy, the genetic improvement of durum wheat was launched in the early 1900s with the exploitation of the genetic variability available in Southern Italy. Moreover, the introduction of exotic landraces from North Africa and West Asia [11] also contributed to the maintenance of a broad genetic variability within Sicilian ecotypes.

Despite the improvements in yield and end-use quality [12,13], modern durum wheat varieties require higher chemical inputs that have a negative impact on the environment, which determined a genetic variability loss due to the replacement of wheat landraces [14,15].

Climate change impacts crop production worldwide, which now requires a large germplasm collection as a source of new traits, especially for bread and durum wheat, which represent the staple food for about 40% of the world's population [16]. Although modern wheat varieties show relevant agronomic performances even in marginal environments, the landraces represent a valuable reservoir of important traits [17–20]. More recently, the evolution of durum wheat from ancient to modern varieties resulted in grain yield and gluten strength increases but a decrease in grain protein concentration. In turn, the quality of pasta improved, but with a decrease in its versatility.

Sicily, the southernmost region of Italy, is the second-largest durum-wheat-producing Italian region, with 343,500 hectares cultivated and 931,800 tons produced in 2021 [5]. Wheat arrived in Sicily around the 3rd millennium B.C. and found an ideal cultivation environment there, but it was only under Roman rule that Sicily became known as the 'Granary of Italy'. Due to its position in the center of the Mediterranean, Sicily is characterized by a subtropical climate that is partially semi-arid, which played a crucial role as a crossroads in the evolutionary history of wheat by acting as a hinge (also in terms of food) between Asia, Africa, and Europe [21]. Over the centuries, a large wheat germplasm collection obtained from the selection of landraces adapted to areas of Sicily has been developed, some of which have been reproduced by farmers. The earliest historical document on Sicilian wheat varieties dates back to the 17th century, but their detailed descriptions proposed in the 20th century mainly are attributable to Ugo De Cillis [22]. 'Timilia' is described as one of the oldest Sicilian varieties [23]; it is currently cultivated throughout Sicily and is also known as 'Tumminia', 'Triminia', 'Diminia', and 'Marzuolo'. It is still used to produce 30% of the local and traditional bread named 'Castelvetro black bread', which is highly appreciated by consumers [24]. De Cillis described the 'Russello' variety, which is spread in the hilly hinterland throughout Sicily except for the western province of Trapani [22].

In addition to these two important local varieties, the Sicilian durum wheat germplasm includes other well-adapted landraces, the names of which have toponymic or morphological origins. Three similar varieties ('Bidi', 'Margherito', and 'Cappelli') originated from a North African population created by Nazzareno Strampelli in 1915.

Rivet wheat (*T. turgidum* subsp. *turgidum*) landrace group or so-called 'Bufale' are distinguished by color, all showing a mealy (softer) fracture of the kernel, a trait closer to bread than durum wheat [25]. Rivet wheat is widespread in the mountain range of the Calabria region, where durum wheat is not largely cultivated [26], and shows rusticity traits that are useful for growing in marginal soils and resistance to diseases [27]

'Farro Lungo' (*T. turgidum* ssp. *turanicum*, synonym 'Perciasacchi') was reported in Sicily at the beginning of the 19th century [28]. Later, De Cillis [22] described it as a typical cultivated wheat in Sicily.

Alongside many durum and tetraploid wheats, a few bread wheat landraces are still cultivated in Sicily; for example, 'Maiorca' and 'Cuccitta' in the northeastern regions and 'Maiorccone' and 'Romano' mainly in the Catania plain.

Recently, due to a greater interest in traditional, healthy, and sustainable food products, an increase in hectareage dedicated to the cultivation of these ancient Sicilian wheat varieties has been observed. To regulate their seed production, a section in the National Variety Register for 'Conservation Varieties' has been established following Directive 98/95/EC [29]; to date, 27 durum wheat and 2 bread wheat Sicilian varieties were registered.

The need for identity conservation of local varieties stimulated research on their genotyping to both decipher the genetic diversity and develop a certification system. In the past, wheat genetic diversity was mostly evaluated using morphological and phenological descriptors [30–32]. More recently, different molecular markers have been utilized, such as single-nucleotide polymorphisms (SNPs) for marker–trait association [33,34], genetic diversity and population structure [35–37], genomic selection [38–40], and quantitative trait loci mapping [41,42]. Several studies focused on Italian wheat germplasm diversity analysis [43–45] using these markers [46–48].

Thanks to the activities of De Cillis and the preservation carried out by Sicilian farmers, a germplasm collection of Sicilian wheat is now conserved in the gene bank of the "Experimental Sicilian Station for wheat crop" (ESS) at Caltagirone (province of Catania, Italy). Over the last century, different populations of the same variety have been distinguished. Currently, the ESS contains more than 200 accessions of different *Triticum* species and subspecies, including durum and common wheat.

This study aimed to assess the morphogenetic diversity of the ESS germplasm collection using a high-density SNP array. In the analysis, durum and bread wheat landraces from the Calabria region were included that could also be conveniently preserved and exploited in future breeding programs. Interestingly, the genetic diversity of the EES collection was compared to the improved varieties of both old and modern wheat using the same SNP array while also defining their population structure. Moreover, the inter- and intrapopulation genetic variations in the historical Sicilian durum wheat varieties were also estimated. This research laid the foundation for expanding the genetic variability of cultivated wheat through the use of these underexplored southern Italian landraces.

2. Results

2.1. Genetic Relationships among *Triticum* Species in the Entire Dataset

The genetic diversity among *Triticum* species, which are widespread in Southern Italy (Sicily and Calabria regions), was investigated by using a high-throughput genotyping system based on a wheat 90k SNP array. After filtering, 20,899 high-quality SNPs out of 81,587 SNPs were retained (Table S1).

A phylogenetic tree on the 126 wheat accessions from Sicily and Calabria was defined via a cluster analysis based on the UPGMA algorithm and the Nei' genetic distance among the SNP-derived genotypes. Four main groups were identified: *T. aestivum*, *T. monococcum*, and two subspecies of *T. turgidum* (Figure 1A).

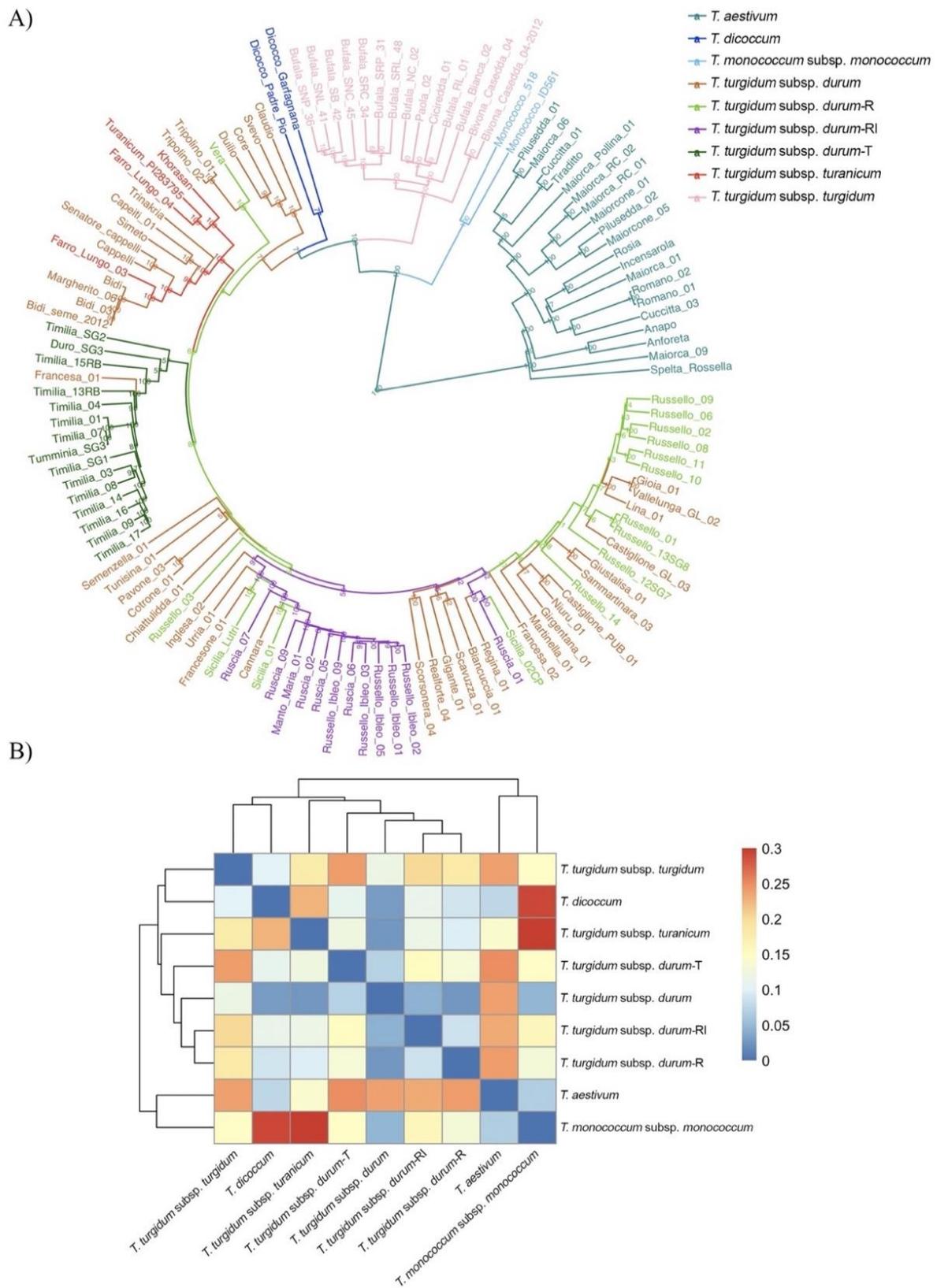


Figure 1. Genetic diversity of 126 wheat genotypes analyzed in the 90k SNP array. (A) UPGMA dendrogram. The varieties are highlighted based on their species and membership group (Table S1). (B) Heatmap based on the population pairwise fixation index (F_{st}) evaluated among the wheat species and groups. R: Russello; RI: Russello Ibleo; T: Timilia.

The first cluster included the *T. aestivum* accessions. Among these, the old varieties from different regions were separated from the modern varieties ‘Anapo’, ‘Anforeta’, and ‘Rossella’ (Table S1). The second cluster was represented by the two *T. monococcum* accessions. The third cluster comprised all the rivet wheat accessions, the ‘Bufale’ group, and three Sicilian old varieties (‘Paola’, ‘Bivona Casedda’, and ‘Ciciredda’). All the durum wheat accessions and the more distant emmer wheat (*T. turgidum* ssp. *dicoccum*) landrace ‘Garfagnana’ and variety ‘Padre Pio’ comprised the fourth-largest cluster. In the latter, it was possible to distinguish several durum wheat subclusters (Figure 1A).

The more recently improved varieties (‘Claudio’, ‘Svevo’, ‘Core’, and ‘Duilio’) were grouped together; the older improved varieties (‘Trinakria’, ‘Senatore Cappelli’, and ‘Capeiti’) were clustered with the older varieties from North Africa (‘Tripolino’, ‘Bidi’, and ‘Margherito’). The turanicum wheat (*T. turgidum* subsp. *turanicum*) gathered within this last subcluster maintained their distinctness except for the “Farro lungo 03” accession. The most mixed subcluster, which encompassed 84 Sicilian durum wheat accessions based on gene bank information, was originally classified into four groups (*durum*-T—‘Timilia’; *durum*-R—‘Russello’; *durum*-RI—‘Russello Ibleo’ or ‘Ruscia’, and *durum*—all the other durum wheat accessions analyzed here). The 14 *durum*-T samples and ‘Francesa_01’ accessions that clustered showed narrow genetic distances. Eleven *durum*-RI accessions were distinguished from the *durum*-R group except for ‘Sicilia Lutri’, Russello03, and Sicilia01. Four landraces from Sicily (‘Urria’, ‘Inglese’, ‘Francesone’, and ‘Cannara’) were also included in this subcluster. Finally, the last subcluster included 11 *durum*-R group, Ruscia01, and 17 durum wheat landraces (*durum* group, Table S2). Among these were ‘Gioia’ and ‘Vallelunga Glabra’ (synonym), ‘Giustalisa’ and ‘Sammaritana’, and ‘Niuru’ and ‘Castiglione Pubescente’ (similar to pairs).

Overall, moderate pairwise fixation index (*Fst*) values among the wheat species and subspecies were recorded (Figure 1B, Table S3). The highest *Fst* value was observed between the diploid *T. monococcum* and the two-tetraploid subsp. *dicoccum* and *turanicum* (0.292 and 0.300, respectively). Instead, a low differentiation (<0.17) was detectable in the other tetraploid subspecies (rivet wheat and ‘Timilia’ and ‘Russello Ibleo’ durum groups), with the lowest value (0.025) highlighted between the pair *T. turgidum* subsp. *durum* and the Russello group (R). Among the four durum wheat groups previously described, a low *Fst* (≤ 0.07) was evident.

The PCoA analysis (Figure S1) clustered the wheat collection into six distinct groups according to the first (PCoA1 \approx 31% of genetic variance) and the second (PCoA2 \approx 10% of genetic variance) principal coordinates. The durum wheat group T and the rivet wheat accessions formed two separate clusters. In contrast, the other durum wheat accessions, including the groups R and RI, as well as the *turanicum* wheat accessions, were gathered in a unique cluster. The population structure of the 126 *Triticum* accessions was inferred using fastSTRUCTURE, which did not account for linkage disequilibrium (LD) between genetic markers (Figure 2). The best number for model complexity (*K*) was between 4 and 6 when using the algorithm for multiple choices *chooseK.py*, with *K* = 6 as the model used to explain the structure in the dataset (Figure S2A). Instead, the prediction error approach recorded the optimum pools at *K* = 5 (Figure S2B). Therefore, the lowest model complexity within the range was explained at this value (*K* = 5). The inferred population membership coefficients are shown in Table S4. A total of 110 accessions (87%) were grouped into 5 gene pools; the remaining 16 genotypes were shown to have an admixture genetic structure using a membership probability threshold of 0.7 (Figure 2, Table S4). The structure analysis confirmed that Francesa01 belonged to the T group (*K*5), in agreement with the phylogenetic (Figure 1A) and PCoA analyses (Figure S1). Among others, five Sicilian accessions (‘Tripolino’, ‘Tunisina’, ‘Chiattulidda’, ‘Pavone’, and ‘Vera’) showed an admixture structure. In addition, Ruscia01 and ‘Duro SG3’, the varieties obtained by the cross Timilia x ‘Biancuccia’ (Table S2), as well as the *dicoccum* ‘Garfagnana’, also showed an admixture structure. All of the bread wheat accessions were included in the same membership group (*K*4), while the rivet wheat accessions were in *K*1 (Table S4). Improved

varieties of durum wheat and the two landraces ‘Bidì’ and ‘Margherito’ were grouped in K2 together with Farro Lungo 03, a Sicilian *turanicum* wheat, and ‘Padre Pio’, a variety obtained from a cross durum wheat x emmer. All of the remaining Sicilian durum wheat accessions, including the R and RI groups, were assigned to cluster K3 (Figure 2, Table S4).

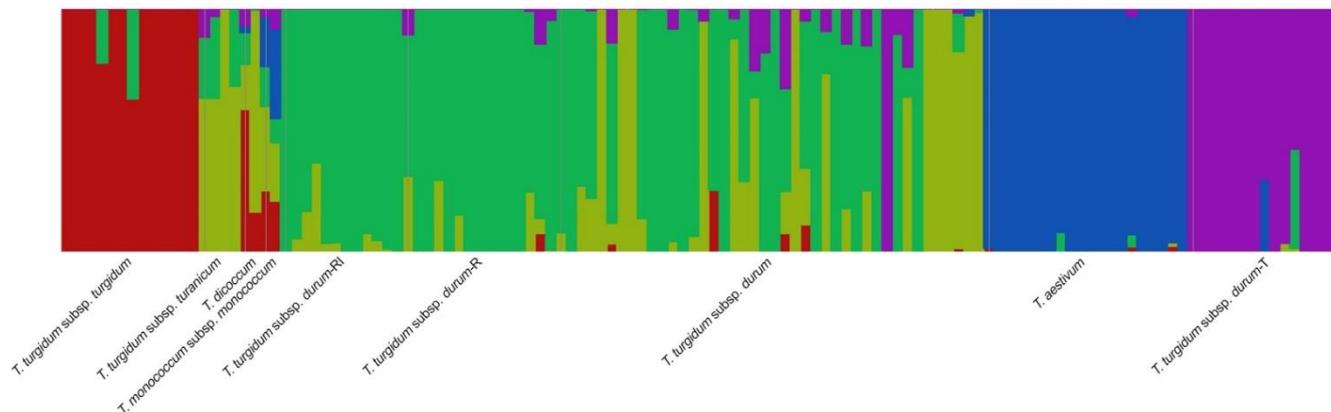


Figure 2. Population structure in the entire wheat dataset (126) as analyzed in the 90k SNP array. The optimum pool number was recorded at $K = 5$ by fastStructure. Each sample is represented by a vertical bar. The color proportion for each bar represents the posterior probability of assignment of each variety to one of the five groups of genetic similarity and ranges from 0 to 100% of the assignment probability range (Table S4).

2.2. Morphological Diversity of Triticum Landraces

Morphological characterization was performed on a subset (110) of the wheat germplasm under study. Sixteen accessions were not included in the analysis based on the twenty-four International Union for the Protection of New Varieties of Plants (UPOV) descriptors (Table S5) because only the germplasm representative of ancient local varieties was considered.

The clustering pattern of the morphological traits was calculated based on the Euclidean distance (Figure 3). All of the bread wheat accessions were grouped into one cluster except for the two Calabrian landraces (‘Rosia’ and ‘Incensarola’), which showed a unique profile that was different from most of the Sicilian accessions, mainly due to plant-height-related traits (var7 and var10) and the absence of glaucosity on the lower leaf blade side (var4). In addition, the presence of awns (var9) was shown by both ‘Maiorca’ from Calabria and two Sicilian ‘Maiorcone’ accessions but was absent in the other Sicilian common wheat accessions. The two Sicilian *T. turgidum* subsp. *turanicum* accessions clustered together with ‘Turanicum_PI283795’, an accession from Afghanistan included in the ESS collection since 2012, and ‘Giustalisa’, a Sicilian landrace of durum wheat.

Another cluster included all seven Calabrian rivet wheat accessions together with BufalaSB02, while the other Sicilian rivet wheat accessions were distributed in different clusters. Six accessions included in the *durum*-R group were clustered together with two durum wheat accessions (VallelungaGL02 and CastiglionePUB01). The remaining accessions of the *durum*-R group were distributed in other clusters together with the durum wheat accessions belonging to the RI and T groups. A slight coloration of the ear characterized the RI and R groups. All of the T group accessions showed a pyramidal ear shape.

A principal component analysis (PCA) of the standardized data was applied to display the wheat morphological trait relationships and their application in germplasm characterization (Figure S3). The first two components explained only 29.1% of the total variation among the morphological traits investigated. However, the first dimension was able to clearly separate the samples belonging to *T. turgidum* subsp. *turanicum* (top-left quadrant) and *T. aestivum* (top-right quadrant), with var3, var5, and var 6 (Table S4) showing the strongest influence ($\cos^2 = 0.60$) (Figure S3). Likewise, almost the entire germplasm of

T. turgidum subsp. *turgidum* studied (11 out of 14 samples; 79%) was included in the bottom-left quadrant. Of the traits, var7 and var10 were the most influential components (Figure S3). Overall, the flag leaf glaucosity of the sheath (var3), ear glaucosity (var6), and plant length (stem, ear, awns, and scurs; var10) were shown to strongly influence PC1 and PC2.

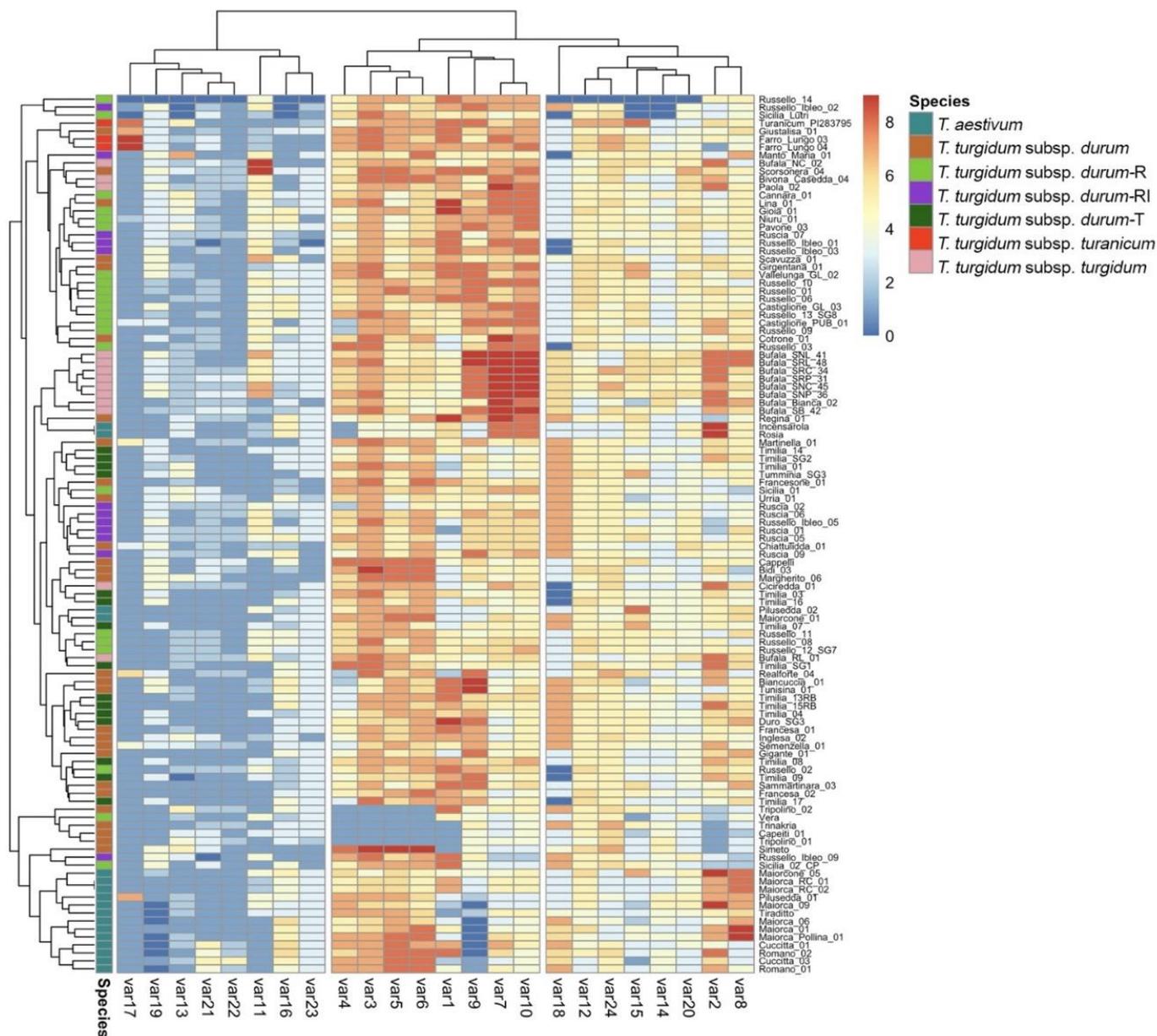


Figure 3. Heatmap of the morphological traits evaluated on a wheat germplasm subset (110) (Table S3). R: ‘Russello’; RI: ‘Russello Ibleo’; T: ‘Timilia’. The colors in the heatmap indicate the values reported in Table S4 for each trait evaluated.

A Pearson correlation matrix showed that variables 3, 4, and 5 (glaucosity traits) were highly correlated in a positive manner, as were var 9 and var10 (ear and plant length) (Figure S4). These were grouped into one of the three main clusters of traits (Figure 3).

2.3. Population Genetics and Diversity of Sicilian Historical Varieties Belonging to Durum Wheat

The genetic background of the samples that belonged only to *T. turgidum* subsp. *durum* (84 accessions) were then evaluated first through an evaluation of their main genetic parameters.

The H_e , H_o , AR, and F_{is} values were calculated (Table 1), which classified the samples into four groups based on the membership assigned after the genetic (Table S3) and morphological evaluation (Table S4). A low H_e was recorded in all four groups that ranged from 0.13 (T) to 0.17 (old durum varieties), as well as a low H_o that ranged from 0.15 (old durum varieties) to 0.20 (RI). Allelic richness (AR) was largely uniform in all four groups, with the highest value recorded by the old durum varieties (Table 1). The F_{is} values, which described the genetic diversity within each group, were negative for all groups except for *T. turgidum* subsp. *durum* (0.130) and ranged from -0.037 (*T. turgidum* subsp. *durum*-R) to -0.301 (*T. turgidum* subsp. *durum*-RI), which highlighted a slight heterozygous excess.

Table 1. Genetic parameters evaluated for the *T. turgidum* subsp. *durum* subset split into four groups: *T. turgidum* subsp. *durum* (all the other accessions), *T. turgidum* subsp. *durum*-R ('Russello'), *T. turgidum* subsp. *durum*-RI ('Russello Ibleo'), and *T. turgidum* subsp. *durum*-T ('Timilia').

Group	H_o	H_e	AR	F_{is}
<i>T. turgidum</i> subsp. <i>durum</i>	0.15	0.17	1.167	0.130
<i>T. turgidum</i> subsp. <i>durum</i> -R	0.16	0.14	1.142	-0.037
<i>T. turgidum</i> subsp. <i>durum</i> -RI	0.20	0.14	1.138	-0.301
<i>T. turgidum</i> subsp. <i>durum</i> -T	0.18	0.13	1.131	-0.199

H_o = Observed heterozygosity; H_e = expected heterozygosity; AR = allelic richness; F_{is} = inbreeding coefficient.

The genetic relationships among the Sicilian durum wheat accessions and the population structure were analyzed using the NJ and DAPC analyses, respectively. Minimum spanning networks (MSNs) were also computed for the same subpanel. The Bayesian information criterion (BIC) method allowed six clusters as the optimal number of groups (Figure S5A). Membership values of each accession to the six clusters are available in the assigned plot (Figure S5B). The DAPC scatter plot showed a triangle-shaped distribution, which highlighted that pool 6 (yellow) was divergent from the others (Figure 4). This pool mainly encompassed accessions belonging to the *T. turgidum* subsp. *durum* (26%, 9 out of 35 genotypes), including the modern durum wheat varieties ('Capeiti', 'Claudio', 'Core', 'Duilio', 'Simeto', 'Svevo', and 'Trinakria'), some very old landraces/varieties ('Tripolino 01' and 'Tripolino 02'), and the 'Vera' accession of the durum-R group (Figure S5B). The other accessions of the *T. turgidum* subsp. *durum* group were mainly split between three pools. These were pool 1 (dark green), which grouped the old varieties 'Bidi', 'Margherito', and 'Senatore Cappelli'; pool 4 (deep pink), which mainly encompassed the accessions of the durum-R group (75%, 15 out of 20 genotypes); and pool 3 (blue), which represented 69% of this pool (11 out of 16 samples). The other two groups—durum-RI and durum-T—were split between two main pools: pool 5 (light green) and pool 2 (orange) (Figure S5B). In pool 5, three durum wheat varieties ('Francesone', 'Inglesa', and 'Urria') and three accessions previously classified in the durum-R group (Sicilia01, 'Sicilia Lutri', and 'Cannara') were also included. Interestingly, the 'DuroSG3' sample, which was selected from a cross between 'Timilia', the old variety 'Biancuccia' belonging to the durum-T group, and the durum-RI accession 'Ruscia 01', were included in Group 3 (Figures 4 and S5B).

Finally, the two accessions 'Castiglione GL03' and 'Castiglione PUB01' together with the old varieties 'Gioia' and 'Niuru' were included in pool 4, while the accession 'Francesa01' was included in pool 2, which was in agreement with the UPGMA cluster analysis. In contrast, the accession 'Francesa02' was assigned to pool 3 together with the old durum wheat varieties (Figure 4).

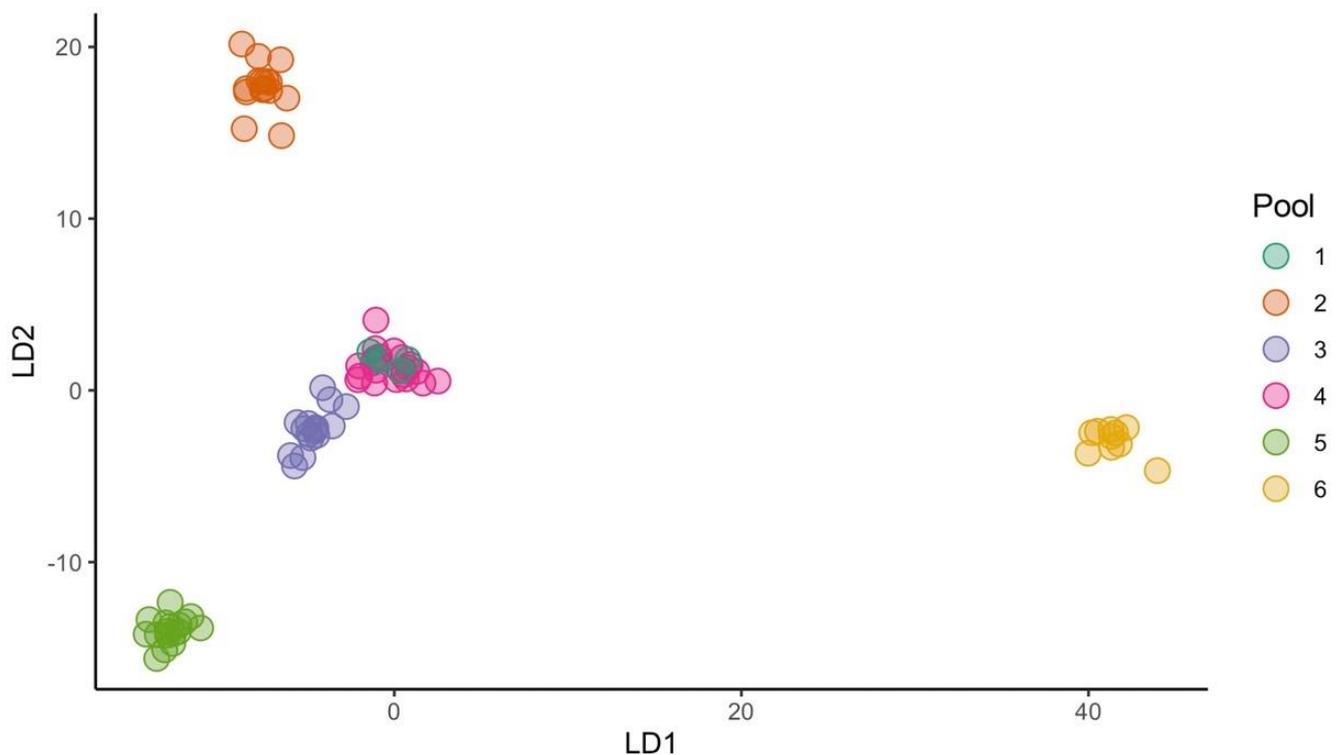


Figure 4. DAPC scatter plot of the 84 accessions belonging to *T. turgidum* subsp. *durum* (Table S1) based on the first and second linear discriminants (LDs) at the best K (6) as evaluated using the BIC method (Figure S5A). The membership for each sample is highlighted in Figure S5B.

The neighbor-joining (NJ) tree, which included only the accessions belonging to *T. turgidum* subsp. *durum*, highlighted a large cluster that gathered all improved durum wheat varieties and the landrace ‘Vera’ (Figure 5A). The remaining accessions belonging to the durum group were split into several clusters that mainly also included genotypes of the durum-R group. Particularly, the durum-R group was largely shifted into two clusters: one shared with ‘Giustalisa’, ‘Lina’, ‘Gigante’, and ‘Sammaritana’, and four old durum wheat varieties, and the other with ‘Girgentana’ and ‘Martinella’ (Figure 5A). In contrast, all accessions of the durum-RI group except Ruscia01 were clustered in a private branch that also included three old Sicilian varieties (‘Francesone01’, ‘Inglesea02’, and ‘Urria01’) and three accessions of the durum-RI group (‘Sicilia02’, ‘Sicilia Lutri’, and ‘Cannara’), which were previously assigned to the durum group, in agreement with the population structure analysis (Figures 4 and S5B). Finally, the durum-T group formed a distinct cluster that included all of the ‘Timilia’ accessions together with the ‘Francesca01’ (in agreement with the DAPC) and ‘Tunisina01’ accessions (Figure 5A).

The minimum spanning network (MSN) confirmed the NJ tree results (Figure 5B), with accessions of the *durum* group (dark green) widespread in the network and admixed mainly with *durum*-R samples, which showed a close relationship (Figure 5B). In contrast, the accessions belonging to *durum*-RI (blue) and *durum*-T (deep pink) appeared in distinct nodes and had a greater diversity than the other groups (Figure 5B).

The multidimensional scaling (MDS) plots (Figure S6) illustrated distinct groups within the durum wheat class. This suggested that a portion of the genetic diversity from the accessions of *durum*-R and *durum*-RI groups had been introgressed in the remaining durum landraces/varieties. In contrast, the *durum*-T group accessions’ gene pool was shown to remain underutilized by breeding.

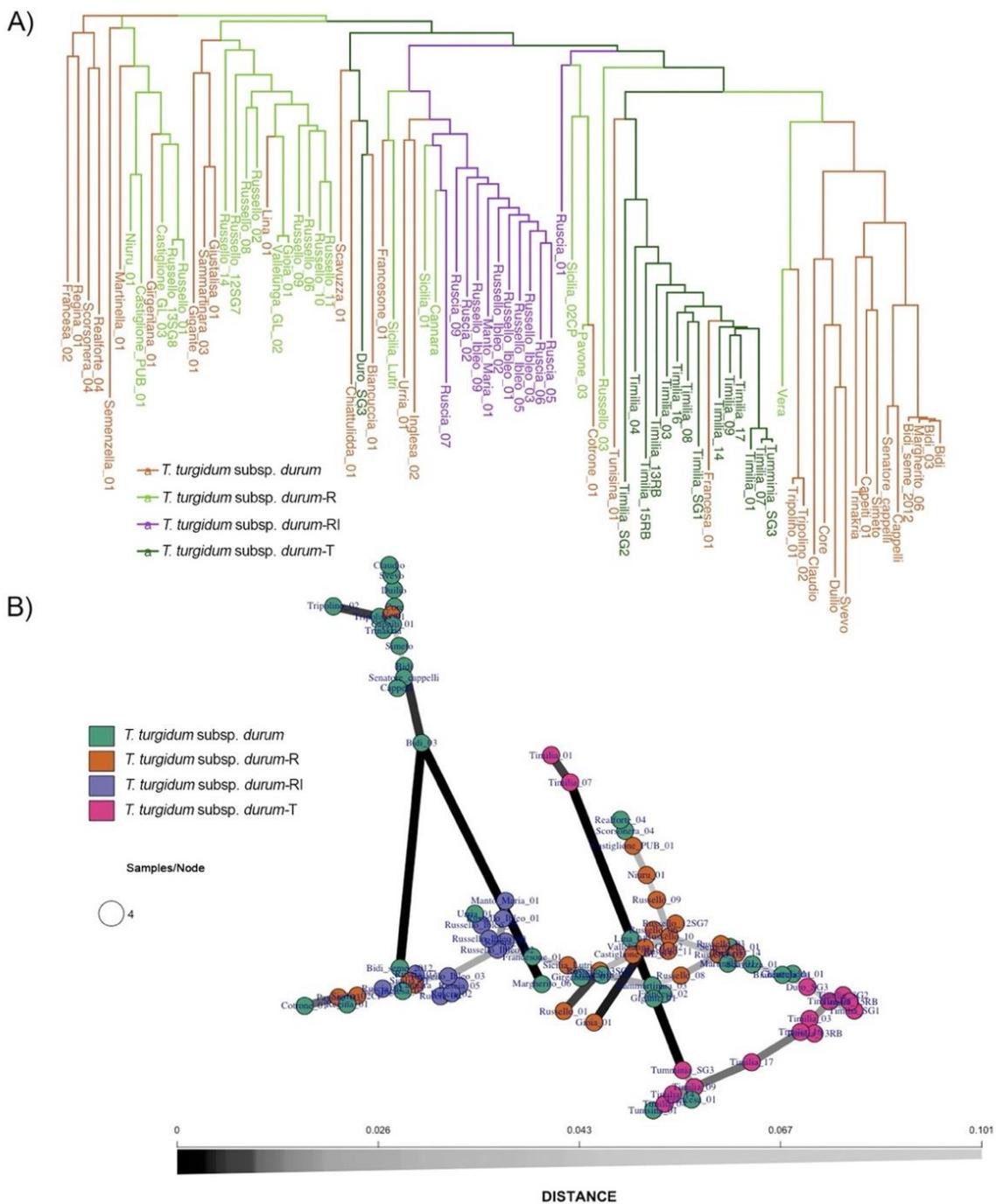


Figure 5. (A) NJ tree based on SNP profiles of varieties belonging to *T. turgidum* subsp. *durum*: *T. turgidum* subsp. *durum*, *T. turgidum* subsp. *durum-R* (‘Russello’), *T. turgidum* subsp. *durum-RI* (‘Russello Ibleo’), and *T. turgidum* subsp. *durum-T* (‘Timilia’). (B) A minimum spanning network (MSN) showing the multilocus genotypes (MLG) of accessions belonging to *T. turgidum* subsp. *durum* subset. The distances between the MLGs indicates Nei’s genetic distance; each MLG is a node and the genetic distance is represented by the edges. The nodes are connected by the minimum distance between samples, which allowed for reticulations. The grey scale of the edges is proportional to the genetic distance; the lines become darker for more-related nodes.

3. Discussion

The presence of wheat in the Mediterranean basin dates to 7000 BP (before the present) [49]; later, wheat spread to North Africa (Tunisia, Morocco, and Algeria) and southern Italy, including Sicily, where it was spread throughout the island. Here, starting

from the 3rd millennium BC, wheat found the best environmental conditions for adapting and evolving. The process of both natural and human selection, as well as the settlement of different populations in Sicily, led to the introduction and/or selection of a large number of local wheat varieties. Therefore, the genetic diversity of *Triticum* landraces cumulated across the centuries and reached a maximum at the beginning of the previous century when breeders began the distribution of new improved varieties [50]. This wide range of genetic variability available in Sicily was described in detail by Ugo De Cillis [22], although the descriptions of some ancient varieties can be found in 17th-century publications [51].

From 1929 until 1948, De Cillis headed the “Experimental Sicilian Station for wheat crop” (ESS) at Caltagirone (province of Catania in Sicily), which collected information on the morphological, physiological, and technological traits of Sicilian wheat landraces/varieties. The first description/classification based on the Percival method [52] placed some limitations on bringing the observed genetic stock into the collection, forcing De Cillis to purify some accessions using recurrent selection and to assign names according to their area of origin. Nowadays, 39 5old varieties of the original core collection of De Cillis’ work have been maintained and incremented by the ESS. This collection represents the large biodiversity available in Sicily since 1900 and is to be preserved as an untapped reservoir for expanding the genetic structure of modern varieties. Furthermore, the growing awareness of the importance of a healthy lifestyle is increasingly linked to the consumption of sustainable alternative foods such as old varieties, which can reduce the environmental burden of human food production. This heightened interest in “ancient grains” was needed to regulate their production and multiplication at the national and regional levels. In turn, this necessitated the development of protocols to avoid commercial fraud and to sustain the economic profits derived from the wheat food chain [29].

Many studies were carried out to characterize and identify wheat species accessions based on morphological characterization according to UPOV traits [53–55], which are, unfortunately, influenced by the environment. In those studies, a relatively large variation based on morphological traits was detected among accessions.

The development of the first molecular marker classes furnished novel tools for avoiding misclassification due to the environment. Among others, the simple sequence repeats (SSR) represent more informative molecular markers that are largely detectable in plant genomes. The genetic diversity of several wheat germplasm collections has been evaluated and distinguished species, subspecies, and genotypes [48,56–58]. Lastly, with the advent of next-generation sequencing (NGS) techniques, a very large panel of SNPs has been set up as a useful tool to provide a detailed reproducible genetic fingerprint of the wheat germplasm [59–61]. SNPs are point mutations that result in single-base-pair divergence among DNA sequences present in both genome coding and noncoding regions. These are also utilized for genetic characterization in wheat, in which the tetraploid and hexaploid conditions made the utilization of microsatellites more difficult.

Several studies assessed the genetic diversity of the Sicilian wheat germplasm described by De Cillis using both SSR [25] and SNP [47,48,62]. Here, we assembled a definitive panel of 126 wheat accessions to assess their genetic relatedness using a large SNP panel. In the collection, bread and durum wheat landraces and varieties (old and modern), *T. monococcum* (EES collection), and many subspecies related to *T. turgidum* (*dicoccum*, *durum*, *turanicum*, and *turgidum*) were included. Our results confirmed the ability of SNP technology to easily distinguish *Triticum* species and subspecies. In addition, we were able to classify 84 durum wheat accessions (including old and modern varieties, as well as landraces) for genetic distances and structures in order to assess their genetic relationships. Moreover, three groups/populations related to ‘Timilia’, ‘Russello’, and ‘Russello Ibleo’ were also included, thereby also defining the intrapopulation genetic variability, as recently reported by Taranto et al. [63].

A study of 56,342 hexaploid accessions from CIMMYT and ICARDA germplasm banks showed that a large portion of genetic variability of hexaploid wheat varieties has not been sampled in breeding selection for modern varieties [64]. In our study, the genetic

distinctness between the ancient bread wheat varieties and two improved varieties ('Anapo' and 'Anforeta') confirmed and suggested a useful gene reserve in the EES collection to be explored and subsequently used in future breeding programs. In the bread wheat cluster, a selected line of the *T. aestivum* subsp. *spelta* and four ancient bread wheat landraces collected in Calabria were also included.

The distinctness between the 14 rivet wheat accessions and the other three tetraploid subspecies, as previously reported [25,62], was confirmed, which was in agreement with De Cillis' description of the landraces 'Bufala Rossa Corta', 'Bufala Rossa Lunga', 'Bufala Nera Corta', 'Bufala Nera Lunga', 'Ciciredda', 'Bivona', and 'Paola' [22]. These underutilized rivet wheat genotypes represent a significant genetic stock that could be useful for maintaining wheat cultivation in specific rural mountainous areas of Sicily and other Mediterranean regions such as Calabria, which are usually considered marginal and unsuitable for elite durum wheat cultivation. In addition, a private cluster related to the rivet wheat accessions (a subcluster from Calabria and another from Sicily) was revealed for the first time, in contrast to Mangini et al. [59]. The UPGMA cluster analysis comprised all the durum wheat accessions in the largest cluster, including four accessions belonging to *T. turgidum* subsp. *turanicum* and, more distant, the two emmer wheat varieties: 'Garfagnana' (a landrace cultivated in Tuscany, Italy) and the variety 'Padre Pio'. This result did not agree with Ganugi et al. [62], who reported three separated clusters for *turanicum*, *dicoccum*, and *durum* accessions.

Khorasan wheat (*T. turgidum* subsp. *turanicum*), better known under the brand Kamut, has recently drawn a large commercial interest due to its high protein content and valuable nutritional qualities [65]. However, there is limited information available on the partitioning of the genetic diversity on the available accessions of Khorasan/*turanicum* wheat. It was reported that they are related to durum wheat [66]. Hence, significant variations in the agronomic traits have been recorded among Khorasan accessions [67]; a univariate analysis was not able to clearly distinguish *turanicum* and *durum* wheat varieties [68], as was also demonstrated in our study. Only the accession Farro lungo 03 showed a posterior probability value close to 1 (0.99998) in the same membership group of durum wheat accessions, while the remaining three accessions, including the 'Khorasan' commercial variety and TuranicumPI283795, showed an admixture genetic structure (K2 and K3).

The genetic analysis was then intensified and limited to the *T. turgidum* subsp. *durum* accessions, which underlined a higher observed heterozygosity in the three R, RI, and T groups compared to the group that included the more recent improved varieties, as could be expected. This observation might infer a mixing of previously isolated populations. The inbreeding coefficient (*F_{is}*) value was close to 1 for the more recent durum wheat varieties (0.130), while the negative values recorded for the T, RI, and R groups highlighted a slight heterozygous excess. Both the 'Timilia' and 'Russello' accessions belonged to the two ancient wheat varieties, which were widely cultivated in Sicily during the 18th and 19th centuries, respectively. Thereafter, the cultivation of these two accessions spread to many Sicilian areas. The accessions belonging to the RI ('Russello Ibleo') group remained confined to their area of origin (Monti Iblei) and appeared to be distinguished by 'Russello' accessions. Recently, intra- and interpopulation genetic variations of two large collections of 'Russello' and 'Timilia' were explored by Taranto et al. [63]. In agreement with our results, a low genetic distance among the seven populations of 'Timilia' was observed, while higher values among the three 'Russello' populations were measured. The same authors identified two different groups within the 'Russello' populations in accordance with their geographical origin: 'Monti Iblei' in southeastern Sicily (named 'Ruscia') and central-western Sicily. Furthermore, the MSN confirmed the UPGMA and NJ tree results for the genetic relationships among the Sicilian durum wheat ancient varieties and *durum*-R samples, while *durum*-RI (blue) and *durum*-T (deep pink) appeared in distinct nodes with greater diversity than other groups. These results agreed with those previously reported by Taranto et al. [46].

Southern European crops after the 2000s have been subjected to high temperatures in summer, early drought, and very hot temperatures at heading, exacerbating the risk of yield loss [69]. Van de Wouw et al. [70,71] reported that genetic erosion in terms of loss of alleles was not observed in wheat before the 1990s. A strengthening of crop resilience can be found within the genetic variability present in landraces and old wheat varieties [44,72,73]. The morphogenetic analysis of the Sicilian *Triticum* germplasm collection maintained at ESS highlighted the importance of preserving this precious reservoir of genetic resources that have developed and evolved over the centuries in the Mediterranean environment.

4. Materials and Methods

4.1. Plant Materials

The germplasm collection included 126 accessions: 2 hexaploid wheats, *T. aestivum* L. (19 accessions), and *T. aestivum* subsp. *spelta* L. Thell. (one); the cultivated einkorn *T. monococcum* L. subsp. *monococcum* (two); four tetraploid subspecies of *T. turgidum* (104) consisting of 14 rivet wheat accessions (seven from Sicily and seven from Calabria); four *T. turgidum* subsp. *turanicum* accessions; two *T. turgidum* subsp. *dicoccum* accessions; 75 Sicilian landraces and nine improved varieties of durum wheat (Table S2). The durum wheat germplasm under study comprised 39 accessions of the original nucleus collected and characterized by De Cillis in 1927, which was partially maintained at the ESS at Caltagirone (province of Catania, Italy). Other accessions were retrieved from the main international germplasm banks (USDA, CGN, IPK, and IHAR). Instead, several accessions of the three historical Sicilian varieties Sicily ('Russello', 'Timilia', and 'Russello Ibleo'/'Ruscia'), which are among the most widespread on the island, have been preserved in the ESS germplasm bank over the last 20 years. All ESS accessions were multiplied in the experimental field of Caltagirone, province of Catania, Sicily (37°05'58" N, 14°29'56" E, 280 m a.s.l.). The germplasm collection from Calabria (seven rivet wheat and four common wheat), which was derived from old local populations, is maintained in the collection at the Department AGRARIA of the University Mediterranea of Reggio Calabria—Italy, in the experimental field of the "Casello" Agricultural Experiment Centre of the Regional Agency for Agriculture (ARSAC), which is located in San Marco Argentano, Calabria (38°10'24" N, 15°45'12" E, 232 m a.s.l.) in a classic cereal-growing area north of Cosenza.

4.2. Bio-Morphological Traits

The bio-morphological characterization included 110 of the 126 genotypes under study. Sixteen genotypes that were not representative of the local varieties from Sicily and Calabria were excluded from this analysis (Table S2).

Twenty-four UPOV bio-morphological traits [74] were measured for each accession (Table S5). The measurements were carried out in the field during the vegetative stages using the Zadoks scale system [75]: from stage 10 (first leaf passing through the coleoptile) to stage 60 (full growth but not flowering) and in the laboratory using a representative sample of 20 ears.

4.3. Statistical Analysis for Morphological and Quality-Related Traits

Samples of wheat grains were analyzed for 24 main morphological traits (Table S5). To easily visualize the differences between samples, a heatmap was developed through R/pheatmap (<https://CRAN.R-project.org/package=pheatmap>; accessed on 13 July 2022). A principal component analysis (PCA) was also performed using the R package FactoMiner [76]. Finally, the Pearson correlation coefficient ($p < 0.05$) was also calculated among the morpho- and quality-related traits using the R/Hmisc package (<https://cran.r-project.org/web/packages/Hmisc/index.html>; accessed on 13 July 2022). A scatter plot with the correlation coefficients and their significance was developed using the PerformanceAnalytics package in R (<https://cran.r-project.org/web/packages/PerformanceAnalytics/index.html>; accessed on 13 July 2022).

4.4. DNA Extraction and SNP Genotyping

Genomic DNA was extracted from fresh leaves of young seedlings of each variety. The GenElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, St. Louis, MO, USA) was used. DNA quantity and quality were checked via both electrophoresis (1% agarose gel) and a NanoDrop®ND-1000 (Thermo Scientific, Waltham, MA, USA).

Two hundred nanograms of genomic DNA were used as template for each reaction in the wheat 90k SNP array [77] for genotyping, following the manufacturer's instructions (Illumina Inc., San Diego, CA, USA). Raw data were visualized and analyzed using the GenomeStudio V2011.1 software (Illumina). The dataset was filtered and standardized as reported in Mercati et al. 2021 [78]. SNPs with an NA rate >1%, a MAF <5%, an NA rate for each individual >20%, and a polymorphic rate between individuals <20% were excluded by subsequent analyses. After filtering, 20,899 high-quality SNPs out of 81,587 SNPs provided were retained (Table S1).

4.5. Cluster, PCoA, and Fst Analysis

Phylogenetic and PCoA analyses were performed to estimate the overall relationship among varieties while considering the entire wheat germplasm investigated (126 genotypes) and the samples only belonging to *T. turgidum* subsp *durum* (84) collected. The distance-based dendrogram with Nei's genetic distance [79], unweighted pair group method with arithmetic mean (UPGMA), neighbor-joining (NJ) algorithm (bootstraps based on 1000 resamplings), and PCoA were developed using R/poppr [80] and R/adegenet [81], respectively, and visualized using R/ggtree [82]. The Wright's population pairwise fixation index (*Fst*) [83] was computed among species and groups using R/HierFstat [84]; a heatmap based on *Fst* values was developed using R/gplots (<https://CRAN.R-project.org/package=gplots>; accessed on 13 July 2022).

4.6. Population Structure Analysis

The number of genetic pools (*K*) was computed for the entire wheat dataset using an admixture model performed through fastSTRUCTURE [85] using the .bed, .bim, and .fam as input files generated by PLINK v1.90 [86] and the structure.py Python script. The algorithm for multiple choices (chooseK.py) was applied to choose the best number for the model complexity (*K*) by evaluating the model complexity that maximized the marginal likelihood and model components used to explain structures in the data. The prediction error for each *K* was also computed using the cross-validation (*-cv*) function to extract the optimum *K*. A Bayesian analysis was also carried out for the *T. turgidum* subsp *durum* subset (84 samples) using a discriminant analysis of principal components (DAPC) implemented in R/adegenet. The Bayesian information criterion (BIC) method [87] was used to infer the *K*-means clustering. The analyses were performed independent of the membership group.

4.7. Population Genetics and Diversity of Sicilian Varieties Belonging to *T. durum*

The filtering procedure and LD (linkage disequilibrium) pruning ($r^2 = 0.20$) resulted in datasets of high-quality SNP markers that were used to evaluate in detail the genetic diversity of samples belonging to *T. turgidum* subsp *durum*. The main genetic parameters (*Ho*, *He*, *Fis*, and allelic richness (AR)) were evaluated using the R/OutFLANK package [88]. A pairwise identical-by-state (IBS) distance among all individuals using R/SNPRelate [89] was calculated and a multidimensional scaling (MDS) analysis was performed. In addition, to evaluate the genetic distances among the varieties gathered in the four groups (Table S1) of *T. turgidum* subsp *durum*, *T. turgidum* subsp *durum*-R (Russello), *T. turgidum* subsp *durum*-RI (Russello Ibleo), and *T. turgidum* subsp *durum*-T (Timilia), a minimum spanning network (MSN) was also developed in R/poppr [80].

5. Conclusions

In conclusion, SNPs represented powerful tools for exploring genetic diversity in *Triticum* germplasm. The relationships among old Southern Italian landraces, popula-

tions, and old and modern varieties were confirmed. We developed a definitive panel of 126 accessions that were mainly preserved at the 'Experimental Sicilian Station (EES) for Granicoltura', including the collection derived from the Nazareno Strampelli research activity and described by De Cillis, among which many landraces and old varieties were cultivated in Sicily and Southern Italy since the beginning of the 20th century. They represent an important reservoir of genes/alleles that can be used in durum and bread wheat breeding to counteract the effects of climate change. The adopted germplasm collection and SNP array allowed us to distinguish not only the *Triticum* species (as was expected), but also many *T. turgidum* subspecies. Interestingly, we were able to distinguish the subspecies *dicoccum* and *turgidum* from most cultivated *durum*. However, the subspecies *turanicum*, including the commercial cultivar Khorasan, was not distinguishable. In addition, two Sicilian old landraces/populations, namely 'Timilia' and 'Russello' and its variant 'Russello Ibleo', were characterized among the 84 durum wheat accessions (landraces and old and modern varieties). The old Sicilian landraces/populations were clearly clustered and distinguished from the remaining durum wheat accessions, including the 'Russello Ibleo' variant. The present research confirmed the importance of the broad genetic base available not only in the Sicilian landraces, but also among the old varieties cultivated in Southern Italy over the last century and recently re-utilized for their high-quality traits. The interpopulation genetic diversity observed for 'Timilia' and 'Russello', together with the synonyms and misclassifications of some old wheat varieties, should be taken into account when registering old varieties in the National Registers of Conservation Varieties. By so doing, it will be possible to avoid commercial fraud and contribute to sustaining the profits of farmers that cultivate landraces and ancient varieties.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232113378/s1>. Figure S1: Principal coordinates analysis (PCoA) of whole wheat germplasm. The varieties are highlighted based on their species and membership group: *T. aestivum* (cyan), *T. dicoccum* (blue), *T. monococcum* subsp. *monococcum* (light blue), *T. turgidum* subsp. *durum* (brown), *T. turgidum* subsp. *durum*-R (Russello, light green), *T. turgidum* subsp. *durum*-RI (Russello Ibleo, violet), *T. turgidum* subsp. *durum*-T (Timilia, dark green), *T. turgidum* subsp. *turanicum* (red), and *T. turgidum* subsp. *turgidum* (pink). PCoA1 \approx 31% of genetic variance; PCoA2 \approx 10% of genetic variance; Figure S2: A) Barplot of the posterior probability for $K = 6$ (the value obtained as the model components used to explain structure in data) evaluated using the algorithm for multiple choices *chooseK.py*. B) The fitting number of model complexity explaining the structure of the entire germplasm investigated (126 varieties) was evaluated for K values ranging from 1 to 9. In addition to the *chooseK.py* algorithm, to extract the optimum K within the range, the prediction error for each K was also computed using the cross validation ($-cv$) function. The lowest model complexity above which prediction errors did not vary significantly was considered the best one, therefore $K = 5$ was chosen; Figure S3: Principal component analysis (PCA) of the morphological traits evaluated for a subset (110) of varieties investigated (Table S3). Samples were organized in seven main groups as follows: (1) *T. aestivum* (cyan), (2) *T. turgidum* subsp. *durum* (brown), (3) *T. turgidum* subsp. *durum*-R (Russello, light green), (4) *T. turgidum* subsp. *durum*-RI (Russello Ibleo, violet), (5) *T. turgidum* subsp. *durum*-T (Timilia, dark green), (6) *T. turgidum* subsp. *turanicum* (red), and (7) *T. turgidum* subsp. *turgidum* (pink). Traits associated ("var"; Table S2) with the discrimination of the samples are indicated in the plot, underlining their significance values ($0.2 < \cos^2 < 0.6$). The medians for each group were also visualized; Figure S4: Pearson correlation matrix of 24 morphological traits. Positive correlations are displayed in blue and negative correlations in red. Color intensities and circle sizes are proportional to the correlation coefficients. Variables with significant correlations ($p < 0.01$) are highlighted in the plot on the right; Figure S5: (A) Bayesian information criterion (BIC) evaluated via the DAPC approach for varieties belonging to *T. turgidum* subsp. *durum*. The BIC approach was used to infer the K-means clustering within the *T. turgidum* subsp. *durum* subset ($K = 6$). (B) Barplot of the posterior probability for the best K (6) of group assignment for each sample belonging to the subset; Figure S6: MDS scatterplot based on the IBS matrix of the *T. turgidum* subsp. *durum* subset. The different groups (Table S4) are highlighted in different colors: *T. turgidum* subsp. *durum* (black), *T. turgidum* subsp. *durum*-R (Russello; red), *T. turgidum* subsp. *durum*-RI (Russello Ibleo; light green), and *T. turgidum* subsp. *durum*-T (Timilia;

light blue); Table S1: SNP profiles (20,899 high-quality SNPs filtered) of samples investigated in the present work (126 varieties). The table is in .raw format. Table S2: List of the 126 accessions belonging to *Triticum* species and subspecies analyzed using SNP markers. For each accession, the accession name, species, passport information and seed sources are reported; Table S3: Pairwise fixation index (*Fst*) values among the wheat species and groups investigated; Table S4: Membership group assigned before and after genetic and morphological evaluation for the entire germplasm collection (126). The probabilities of posterior membership for each sample included in the present study are reported. The varieties with a new membership group after genetic and morphological evaluation are underlined. The varieties belonging to the *T. turgidum* subsp. *durum* subset (84) used for the following analysis are in bold. R: Russello, RI: Russello Ibleo, T: Timilia; Table S5: Morphological traits evaluated for a subset (110) of varieties investigated. In the table, the species/group for each variety is highlighted. The code (“var”) reported in Figures 3 and S3 for each trait is also included. R: Russello, RI: Russello Ibleo, T: Timilia.

Author Contributions: Conceptualization, M.C.F., F.M. and F.S.; formal analysis, M.C.F., F.M. and G.P. (Guglielmo Puccio); investigation, M.C.F., F.M., A.S., G.P. (Giovanni Preiti), A.M. and S.B. (Sara Bosi); resources, S.B. (Sara Bosi), F.M., M.C.F. and I.M.; data curation, M.C.F., F.M., A.S., G.P. (Giovanni Preiti), and S.B. (Sebastiano Blangiforti); writing—original draft preparation, M.C.F., F.M., A.S., G.P. (Giovanni Preiti), S.B. (Sebastiano Blangiforti), and F.S.; writing—review and editing, M.C.F., F.M. and F.S.; visualization, M.C.F. and F.M.; supervision, M.C.F. and F.M.; funding acquisition, S.B. (Sara Bosi). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project “SAVE” (Protection and promotion of ancient common wheat varieties in the Emilia Romagna region) under grant number 5015855. The project was supported by “Regione Emilia Romagna” under the “PSR 2014–2020” program, Focus Area 4A.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the text or Supplementary Material here.

Acknowledgments: The authors wish to thank Anne Whittaker for scientific English language editorial assistance and Maria Teresa Sardina of the University of Palermo for her support in the laboratory activities.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Poole, N.; Donovan, J.; Erenstein, O. Viewpoint: Agri-Nutrition Research: Revisiting the contribution of maize and wheat to human nutrition and health. *Food Policy* **2021**, *100*, 101976. [CrossRef] [PubMed]
2. Shewry, P.R.; Hey, S.J. The Contribution of Wheat to Human Diet and Health. *Food Energy Secur.* **2015**, *4*, 178–202. [CrossRef] [PubMed]
3. FAOSTAT. Available online: <http://www.fao.org/faostat> (accessed on 24 May 2022).
4. Guzmán, C.; Autrique, J.E.; Mondal, S.; Singh, R.P.; Govindan, V.; Morales-Dorantes, A.; Posadas-Romano, G.; Crossa, J.; Ammar, K.; Peña, R.J. Response to drought and heat stress on wheat quality, with special emphasis on bread-making quality, in durum wheat. *Field Crop. Res.* **2016**, *186*, 157–165. [CrossRef]
5. ISTAT. Available online: http://dati.istat.it/Index.aspx?DataSetCode=DCSP_COLTIVAZIONI (accessed on 24 May 2022).
6. McFadden, E.S.; Sears, E.R. The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.* **1946**, *37*, 81–89. [CrossRef] [PubMed]
7. Martínez-Moreno, F.; Solís, I.; Noguero, D.; Blanco, A.; Özberk, İ.; Nsarellah, N.; Elias, E.; Mylonas, I.; Soriano, J.M. Durum wheat in the mediterranean rim: Historical evolution and genetic resources. *Genet. Resour. Crop Evol.* **2020**, *67*, 1415–1436. [CrossRef]
8. Petersen, G.; Seberg, O.; Yde, M.; Berthelsen, K. Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Mol. Phylogenet. Evol.* **2006**, *39*, 70–82. [CrossRef]
9. de Sousa, T.; Ribeiro, M.; Sabeça, C.; Igrejas, G. The 10,000-year success story of wheat! *Foods* **2021**, *10*, 2124. [CrossRef]
10. Tanksley, S.D.; McCouch, S.R. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* **1997**, *277*, 1063–1066. [CrossRef]
11. Scarascia Mugnozza, G.T. The contribution of Italian wheat geneticists: From Nazareno Strampelli to Francesco D’ Amato. *Rome Accad. Naz. Delle Sci.* **2005**, *2005*, 53–75.

12. Álvaro, F.; Isidro, J.; Villegas, D.; García del Moral, L.F.; Royo, C. Breeding effects on grain filling, biomass partitioning, and remobilization in mediterranean durum wheat. *Agron. J.* **2008**, *100*, 361–370. [[CrossRef](#)]
13. Spina, A.; Dinelli, G.; Palumbo, M.; Whittaker, A.; Cambrea, M.; Negri, L.; Bosi, S. Evaluation of standard physico-chemical and rheological parameters in predicting bread-making quality of durum wheat (*Triticum Turgidum* L. ssp. *durum* [Desf.] Husn.). *Int. J. Food Sci. Technol.* **2021**, *56*, 3278–3288. [[CrossRef](#)]
14. Cappelli, A.; Cini, E. Challenges and opportunities in wheat flour, pasta, bread, and bakery product production chains: A systematic review of innovations and improvement strategies to increase sustainability, productivity, and product quality. *Sustainability* **2021**, *13*, 2608. [[CrossRef](#)]
15. Brown, A.H.D. Variation under domestication in plants: 1859 and today. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 2523–2530. [[CrossRef](#)] [[PubMed](#)]
16. FAO. *How to Feed the World—2050: High-Level Expert Forum*; FAO: Rome, Italy, 2009.
17. De Vita, P.; Nicosia, O.L.D.; Nigro, F.; Platani, C.; Riefolo, C.; Di Fonzo, N.; Cattivelli, L. Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *Eur. J. Agron.* **2007**, *26*, 39–53. [[CrossRef](#)]
18. Dwivedi, S.L.; Ceccarelli, S.; Blair, M.W.; Upadhyaya, H.D.; Are, A.K.; Ortiz, R. Landrace germplasm for improving yield and abiotic stress adaptation. *Trends Plant Sci.* **2016**, *21*, 31–42. [[CrossRef](#)]
19. Lopes, M.S.; El-Basyoni, I.; Baenziger, P.S.; Singh, S.; Royo, C.; Ozbek, K.; Aktas, H.; Ozer, E.; Ozdemir, F.; Manickavelu, A.; et al. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *J. Exp. Bot.* **2015**, *66*, 3477–3486. [[CrossRef](#)] [[PubMed](#)]
20. Landi, S.; Hausman, J.-F.; Guerriero, G.; Esposito, S. *Poaceae* vs. abiotic stress: Focus on drought and salt stress, recent insights and perspectives. *Front. Plant Sci.* **2017**, *8*, 1214. [[CrossRef](#)]
21. Venora, G.; Blangiforti, S. *I Grani Antichi Siciliani*; Le Fate, Ed.; 2017; ISBN 978-88-94096-6-5.
22. De Cillis, U. *I Frumenti Siciliani*; Tipografia Zuccarello & Izzi: Catania, Italy, 1942; ISBN 88-7751-229-6.
23. Theophrastus. Enquiry into Plants, Volume II: Books 6-9. In *On Odours. Weather Signs*; Hort, A.F., Ed.; Harvard University Press: Cambridge, MA, USA, 1916.
24. Giancaspro, A.; Colasuonno, P.; Zito, A.; Blanco, A.; Pasqualone, A.; Gadaleta, A. Varietal traceability of bread 'Pane Nero di Castelvetro' by denaturing high pressure liquid chromatography analysis of single nucleotide polymorphisms. *Food. Control.* **2016**, *59*, 809–817. [[CrossRef](#)]
25. Sicilia, A.; Anastasi, U.; Bizzini, M.; Montemagno, S.; Nicotra, C.; Blangiforti, S.; Spina, A.; Cosentino, S.L.; Lo Piero, A.R. Genetic and morpho-agronomic characterization of sicilian tetraploid wheat germplasm. *Plants* **2022**, *11*, 130. [[CrossRef](#)]
26. Dubcovsky, J.; Dvorak, J. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **2007**, *316*, 1862–1866. [[CrossRef](#)]
27. Quaranta, F.; Belocchi, A.; Fornara, M.; Ripa, C.; D'Egidio, M.G. Le Varietà di Frumento duro in Italia: Risultati della rete Nazionale di Sperimentazione 1999–2012. In Proceedings of the Council for Agricultural Research and Economics, Rome, Italy, 2013; ISBN 978-88-97081-31-9.
28. Balsamo, P. *Giornale del Viaggio Fatto in Sicilia e Particolarmente Nella Contea di Modica*; Reale Stamperia: Palermo, Italy, 1809; Volume 15.
29. Legislative Decree n. 149, of the 29/10/2009 GU n. 254. 31 October 2009. Available online: <https://www.fao.org/faolex/results/details/en/c/LEX-FAOC090693> (accessed on 16 August 2022).
30. Aghaee, M.; Mohammadi, R.; Nabovati, S. Agro-morphological characterization of durum wheat accessions using pattern analysis. *Aust. J. Crop Sci.* **2010**, *4*, 505–514.
31. Gurcan, K.; Demirel, F.; Tekin, M.; Demirel, S.; Akar, T. Molecular and agro-morphological characterization of ancient wheat landraces of turkey. *BMC Plant. Biol.* **2017**, *17*, 171. [[CrossRef](#)] [[PubMed](#)]
32. Gharib, M.A.A.H.; Qabil, N.; Salem, A.H.; Ali, M.M.A.; Awaad, H.A.; Mansour, E. Characterization of wheat landraces and commercial cultivars based on morpho-phenological and agronomic traits. *Cereal Res. Commun.* **2021**, *49*, 149–159. [[CrossRef](#)]
33. Wang, S.X.; Zhu, Y.L.; Zhang, D.X.; Shao, H.; Liu, P.; Hu, J.B.; Zhang, H.; Zhang, H.P.; Chang, C.; Lu, J.; et al. Genome-wide association study for grain yield and related traits in elite wheat varieties and advanced lines using SNP markers. *PLoS ONE* **2017**, *12*, e0188662. [[CrossRef](#)] [[PubMed](#)]
34. Mourad, A.M.; Sallam, A.; Belamkar, V.; Wegulo, S.; Bowden, R.; Jin, Y.; Mahdy, E.; Bakheit, B.; El-Wafaa, A.A.; Poland, J.; et al. Genome-wide association study for identification and validation of novel snp markers for *Sr6* stem rust resistance gene in bread wheat. *Front. Plant Sci.* **2018**, *9*, 380. [[CrossRef](#)]
35. 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*). *BMC Genet.* **2020**, *21*, 18. [[CrossRef](#)]
36. Rufo, R.; Alvaro, F.; Royo, C.; Soriano, J.M. From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers. *PLoS ONE* **2019**, *14*, e0219867. [[CrossRef](#)] [[PubMed](#)]
37. Kumar, D.; Chhokar, V.; Sheoran, S.; Singh, R.; Sharma, P.; Jaiswal, S.; Iquebal, M.A.; Jaiswar, A.; Jaisri, J.; Angadi, U.B.; et al. Characterization of genetic diversity and population structure in wheat using array based SNP markers. *Mol. Biol. Rep.* **2020**, *47*, 293–306. [[CrossRef](#)]

38. Poland, J.; Endelman, J.; Dawson, J.; Rutkoski, J.; Wu, S.; Manes, Y.; Dreisigacker, S.; Crossa, J.; Sánchez-Villeda, H.; Sorrells, M.; et al. Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* **2012**, *5*, 103–113. [[CrossRef](#)]
39. Belamkar, V.; Guttieri, M.J.; Hussain, W.; Jarquín, D.; El-basyoni, I.; Poland, J.; Lorenz, A.J.; Baenziger, P.S. Genomic selection in preliminary yield trials in a winter wheat breeding program. *G3 Genes Genomes Genet.* **2018**, *8*, 2735–2747. [[CrossRef](#)]
40. Larkin, D.L.; Lozada, D.N.; Mason, R.E. Genomic Selection—Considerations for successful implementation in wheat breeding programs. *Agronomy* **2019**, *9*, 479. [[CrossRef](#)]
41. Assanga, S.O.; Fuentealba, M.; Zhang, G.; Tan, C.; Dhakal, S.; Rudd, J.C.; Ibrahim, A.M.H.; Xue, Q.; Haley, S.; Chen, J.; et al. Mapping of quantitative trait loci for grain yield and its components in a US popular winter wheat TAM 111 using 90K SNPs. *PLoS ONE* **2017**, *12*, e0189669. [[CrossRef](#)] [[PubMed](#)]
42. Ren, T.; Fan, T.; Chen, S.; Chen, Y.; Ou, X.; Jiang, Q.; Peng, W.; Ren, Z.; Tan, F.; Luo, P.; et al. Identification and validation of quantitative trait loci for the functional stay green trait in common wheat (*Triticum aestivum* L.) via high-density SNP-based genotyping. *Theor. Appl. Genet.* **2022**, *135*, 1429–1441. [[CrossRef](#)] [[PubMed](#)]
43. Martos, V.; Royo, C.; Rharrabti, Y.; Garcia Del Moral, L.F. Using AFLPs to determine phylogenetic relationships and genetic erosion in durum wheat cultivars released in Italy and Spain throughout the 20th century. *F. Crop. Res.* **2005**, *91*, 107–116. [[CrossRef](#)]
44. Figliuolo, G.; Mazzeo, M.; Greco, I. Temporal variation of diversity in Italian durum wheat germplasm. *Genet. Resour. Crop Evol.* **2007**, *54*, 615–626. [[CrossRef](#)]
45. Ormoli, L.; Costa, C.; Negri, S.; Perenzin, M.; Vaccino, P. Diversity trends in bread wheat in Italy during the 20th century assessed by traditional and multivariate approaches. *Sci. Rep.* **2015**, *5*, 8574. [[CrossRef](#)]
46. Taranto, F.; D’Agostino, N.; Rodriguez, M.; Pavan, S.; Minervini, A.P.; Pecchioni, N.; Papa, R.; De Vita, P. Whole genome scan reveals molecular signatures of divergence and selection related to important traits in durum wheat germplasm. *Front. Genet.* **2020**, *11*, 217. [[CrossRef](#)]
47. Mangini, G.; Margiotta, B.; Marcotuli, I.; Signorile, M.A.; Gadaleta, A.; Blanco, A. Genetic diversity and phenetic analysis in wheat (*Triticum turgidum* subsp. *durum* and *Triticum aestivum* subsp. *aestivum*) landraces based on SNP markers. *Genet. Resour. Crop Evol.* **2017**, *64*, 1269–1280. [[CrossRef](#)]
48. Fiore, M.C.; Mercati, F.; Spina, A.; Blangiforti, S.; Venora, G.; Dell’Acqua, M.; Lupini, A.; Preiti, G.; Monti, M.; Pè, M.E.; et al. High-throughput genotype, morphology, and quality traits evaluation for the assessment of genetic diversity of wheat landraces from sicily. *Plants* **2019**, *8*, 116. [[CrossRef](#)]
49. Feldman, M. Origin of Cultivated Wheat. In *The World Wheat Book: A History of Wheat Breeding*; Bonjean, A.P., Angus, W.J., Eds.; Tec & Doc/Intercept Ltd.: Paris, France, 2001; pp. 3–56.
50. Perrino, P.; Hammer, K. Sicilian wheat varieties. *Die Kulturpflanze* **1983**, *31*, 227–279. [[CrossRef](#)]
51. Mazzola, P.; Raimondo, F.M.; Schicchi, R. The agro-biodiversity of Sicily in ancient herbaria and illustrated works. *Bocconea* **2003**, *16*, 311–321.
52. Percival, J. *The Wheat Plant*; Duckworth and Co.: London, UK, 1921; p. 628. [[CrossRef](#)]
53. Pecetti, L.; Doust, M.A.; Calcagno, L.; Raciti, C.N.; Boggini, G. Variation of morphological and agronomical traits, and protein composition in durum wheat germplasm from eastern Europe. *Genet. Resour. Crop Evol.* **2001**, *48*, 609–620. [[CrossRef](#)]
54. Sciacca, F.; Cambrea, M.; Licciardello, S.; Pesce, A.; Romano, E.; Spina, A.; Virzi, N.; Palumbo, M. Evolution of Durum Wheat from Sicilian Landraces to Improved Varieties. In *Options Méditerranéennes: Série A. Séminaires Méditerranéens, Proceedings of the International Symposium on Genetics and Breeding of Durum Wheat, Rome, Italy, 27–30 May 2013*; Porceddu, E., Damania, A.B., Qualset, C.O., Eds.; Centre International de Hautes Etudes Agronomiques Méditerranéennes: Montpellier, France, 2014; pp. 139–145.
55. Ouaja, M.; Bahri, B.A.; Aouini, L.; Ferjaoui, S.; Medini, M.; Marcel, T.C.; Hamza, S. Morphological characterization and genetic diversity analysis of Tunisian durum wheat (*Triticum turgidum* var. *durum*) accessions. *BMC Genom.* **2021**, *22*, 3. [[CrossRef](#)] [[PubMed](#)]
56. Senturk Akfirat, F.; Uncuoglu, A.A. Genetic Diversity of Winter Wheat (*Triticum aestivum* L.) Revealed by SSR Markers. *Biochem. Genet.* **2013**, *51*, 223–229. [[CrossRef](#)]
57. Sajjad, M.; Khan, S.H.; Shahzad, M. Patterns of allelic diversity in spring wheat populations by SSR-markers. *Cytol. Genet.* **2018**, *52*, 155–160. [[CrossRef](#)]
58. Marzario, S.; Logozzo, G.; David, J.L.; Zeuli, P.S.; 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. [[CrossRef](#)] [[PubMed](#)]
59. Mangini, G.; Nigro, D.; Margiotta, B.; De Vita, P.; Gadaleta, A.; Simeone, R.; Blanco, A. Exploring SNP diversity in wheat landraces germplasm and setting of a molecular barcode for fingerprinting. *Genet. Resour. Crop Evol.* **2018**, *46*, 377–387. [[CrossRef](#)]
60. Gao, L.; Jia, J.; Kong, X. A SNP-based molecular barcode for characterization of common wheat. *PLoS ONE* **2016**, *11*, e0150947. [[CrossRef](#)]
61. Dreisigacker, S.; Sharma, R.K.; Huttner, E.; Karimov, A.; Obaidi, M.Q.; Singh, P.K.; Sansaloni, C.; Shrestha, R.; Sonder, K.; Braun, H.J. Tracking the adoption of bread wheat varieties in Afghanistan using DNA fingerprinting. *BMC Genomics* **2019**, *20*, 660. [[CrossRef](#)]

62. Ganugi, P.; Masoni, A.; Sbrana, C.; Dell'Acqua, M.; Pietramellara, G.; Benedettelli, S.; Avio, L.M. Genetic variability assessment of 127 *Triticum turgidum* L. accessions for mycorrhizal susceptibility-related traits detection. *Sci. Rep.* **2021**, *11*, 13426. [CrossRef]
63. Taranto, F.; Di Serio, E.; Miazzi, M.M.; Pavan, S.; Saia, S.; De Vita, P.; D'Agostino, N. Intra and inter-population genetic diversity of "Russello" and "Timilia" landraces from Sicily: A proxy towards the identification of favorable alleles in durum wheat. *Agronomy* **2022**, *12*, 1326. [CrossRef]
64. Sansaloni, C.; Franco, J.; Santos, B.; Percival-Alwyn, L.; Singh, S.; Petroli, C.; Campos, J.; Dreher, K.; Pixley, K. Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nat. Commun.* **2020**, *11*, 4572. [CrossRef] [PubMed]
65. Sissons, M.J.; Hare, R.A. Tetraploid wheat—A resource for genetic improvement of durum wheat quality. *Cereal Chem.* **2002**, *79*, 78–84. [CrossRef]
66. Mac Key, J. Wheat: Its Concept, Evolution and Taxonomy. In *Durum Wheat Breeding: Current Approaches and Future Strategies*; Royo, C., Nachit, M.M., Di Fonzo, N., Araus, J.L., Pfeiffer, W.H., Slafer, G.A., Eds.; Food Products Press: Binghamton, NY, USA, 2005; Volume 1, pp. 3–61.
67. Iannucci, A.; Codianni, P. Phenotypic parent selection within a khorasan wheat collection and genetic variation in advanced breeding lines derived by hybridization with durum wheat. *Front. Plant Sci.* **2019**, *10*, 1460. [CrossRef] [PubMed]
68. Tavoletti, S.; Pasquini, M.; Mozzon, M.; Servadio, D.; Merletti, A.; Mannozi, C.; Foligni, R. Discrimination among varieties of *Triticum turgidum* subspecies (*dicocon*, *turanicum* and *durum*) based on the fatty acid profile. *J. Cereal Sci.* **2021**, *99*, 103213. [CrossRef]
69. Semenov, M.A.; Shewry, P.R. Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe. *Sci. Rep.* **2011**, *1*, 66. [CrossRef]
70. Van de Wouw, M.; Kik, C.; Van Hintum, T.; Van Treuren, R.; Visser, B. Genetic erosion in crops: Concept, research results and challenges. *Plant Genet. Res.* **2010**, *8*, 1–15. [CrossRef]
71. Van de Wouw, M.; Van Hintum, T.; Kik, C.; Van Treuren, R.; Visser, B. Genetic diversity trends in twentieth century crop cultivars: A meta analysis. *Theor. Appl. Genet.* **2010**, *120*, 1241–1252. [CrossRef]
72. Kahiluoto, H.; Kaseva, J.; Balek, J.; Olesen, J.E.; Ruiz-Ramos, M.; Gobin, A.; Kersebaum, K.C.; Takáč, J.; Ruget, F.; Ferrise, R.; et al. Decline in climate resilience of European wheat. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 123–128. [CrossRef]
73. Kyratzis, A.C.; Nikoloudakis, N.; Katsiotis, A. Genetic variability in landraces populations and the risk to lose genetic variation. The example of landrace 'Kyperounda' and its implications for ex situ conservation. *PLoS ONE* **2019**, *14*, e0224255. [CrossRef]
74. International Union for the Protection of New Varieties of Plants (UPVO) Wheat and Durum Wheat. 2012. Available online: <https://www.upov.int> (accessed on 20 May 2022).
75. Zadoks, J.C.; Chang, T.T.; Konzak, C.F. A decimal code for the growth stages of cereals. *Weed Res.* **1974**, *14*, 415–421. [CrossRef]
76. Husson, F.; Josse, J.; Lê, S. FactoMineR: An R Package for multivariate analysis. *J. Stat. Soft.* **2008**, *25*, 1–18.
77. Wang, S.; Wong, D.; Forrest, K.; Allen, A.; Chao, S.; Huang, B.E.; Maccaferri, M.; Salvi, S.; Milner, S.G.; Cattivelli, L.; et al. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotech. J.* **2014**, *12*, 787–796. [CrossRef] [PubMed]
78. Mercati, F.; De Lorenzis, G.; Mauceri, A.; Zerbo, M.; Brancadoro, L.; D'Onofrio, C.; Morcia, C.; Barbagallo, M.G.; Bignami, C.; Gardiman, M.; et al. Integrated bayesian approaches shed light on the dissemination routes of the eurasian grapevine germplasm. *Front. Plant Sci.* **2021**, *12*, 692661. [CrossRef]
79. Nei, M. Genetic distance between populations. *Am. Nat.* **1972**, *106*, 283–292. [CrossRef]
80. Kamvar, Z.N.; Brooks, J.C.; Grünwald, N.J. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front. Genet.* **2015**, *6*, 208. [CrossRef] [PubMed]
81. Jombart, T. ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [CrossRef] [PubMed]
82. Yu, G. Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinform.* **2020**, *69*, e96. [CrossRef]
83. Wright, T. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **1965**, *19*, 395–420. [CrossRef]
84. Goudet, J. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **2005**, *5*, 184–186. [CrossRef]
85. Raj, A.; Stephens, M.; Pritchard, J.K. fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics* **2014**, *197*, 573–589. [CrossRef] [PubMed]
86. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef] [PubMed]
87. Neath, A.A.; Cavanaugh, J.E. The Bayesian information criterion: Background, derivation, and applications. *WIREs Comput. Stat.* **2012**, *4*, 199–203. [CrossRef]
88. Privé, F.; Aschard, H.; Blum, M.G.B. Efficient management and analysis of large-scale genome-wide data with two R packages: Bigstatsr and bigsnpr. *bioRxiv* **2017**. bioRxiv:190926. [CrossRef]
89. Zheng, X.; Levine, D.; Shen, J.; Gogarten, S.M.; Laurie, C.; Weir, B.S. A High-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **2012**, *28*, 3326–3328. [CrossRef] [PubMed]