Food Bioscience 79 (2026) 108750

Contents lists available at ScienceDirect

Food Bioscience

journal homepage: www.elsevier.com/locate/fbio

Assessing genetic diversity and developing traceability tools for the Sicilian tetraploid wheat conservation varieties supply chain

Marco Bonarrigo [a], Giulio Metelli [a], Bernardo Messina [b], Giuseppe Russo [b], Francesco Sestili [a], Stefania Masci [a,*]

[a] Department of Agriculture and Forest Science (DAFNE), University of Tuscia, Via San Camillo de Lellis Snc, 01100, Viterbo, VT, Italy
[b] Consorzio di Ricerca Gian Pietro Ballatore, Z.I. Dittaino, 94010, Assoro, EN, Italy

ARTICLE INFO

Keywords:
Durum wheat
Sicilian conservation varieties
SNP genotyping
Traceability
KASP

ABSTRACT

Durum wheat landraces and historical varieties, commonly referred to as conservation varieties, are gaining an increasing interest lately, due to their agronomic, cultural, and economic significance. In Sicily, these varieties are still cultivated and play a crucial role in supporting local agricultural systems. To promote their value and ensure product authenticity along the supply chain, reliable traceability tools are needed. We performed SNP genotyping on a collection of 181 tetraploid wheat accessions using the Illumina Infinium 7K SNP array. The panel included 141 accessions representing six key Sicilian conservation varieties (Bidì, Perciasacchi, Ruscìa, Russello, Timilia reste bianche, and Timilia reste nere) and 40 additional accessions comprising Italian landraces, historical cultivars, and modern varieties used as references. Of these, 139 accessions belonged to *Triticum turgidum* ssp. *durum*, while 42 were identified as *T. turgidum* ssp. *turanicum*. Clustering analyses based on multi-locus genotypes and identity-by-state distances revealed several misidentified accessions, likely due to seed lot contamination maintained at the farm level. Notably, Russello, Ruscìa, and Timilia showed marked genetic divergence from both traditional and modern cultivars, underscoring their distinct genetic identities. In contrast, Turanicum accessions displayed low genetic differentiation, suggesting a shared origin and limited diversification. From the overall dataset, 18 SNP markers were identified as highly effective for distinguishing Sicilian conservation varieties. Kompetitive Allele-Specific PCR (KASP) assays were developed and validated for these markers, providing the foundation for a robust, DNA-based traceability system applicable throughout the conservation variety supply chain.

1. Introduction

Durum wheat (*Triticum turgidum* L. ssp. *durum* (Desf.) Husn., 2n = 4x = 28 AABB genome) is an important crop worldwide, with an annual production of about 37 million tonnes (Bożek et al., 2021). It is particularly cultivated in the Mediterranean area, where it is typically used for production of pasta, couscous, bulgur and some types of bread, but its cultivation extends also to Northern Europe Countries such as Denmark, Finland and Sweden. This crop is especially important in Italy, which is in fact one of the largest producers in the world and the first country for pasta production (Blanco, 2024).

The historical importance of this crop for Italy is evidenced by the fact that the first breeding programs started there, at the beginning of the 1900s. In that period, durum wheat cultivation was based mainly on landraces adapted and evolved in the southern regions of Italy, especially in Sicily. First breeding programs involved the selection of pure lines from these landraces and from landraces originating from North Africa. The most famous cultivar developed in this way was Cappelli (*syn.* Senatore Cappelli), selected by Nazareno Strampelli from the Tunisian landrace Jean Retifah and released in 1915, although E. De Cillis (1927) and some recent molecular studies suggest that it was selected from Bidì or Mahmoudi, two other Tunisian landraces (Fiore et al., 2019; Marzario et al., 2018; Taranto et al., 2020). The selection process brought the release of different varieties, most belonging to the *mediterraneum* type, until the mid of the 19th century, when breeders started to cross them with the *syriacum* type landraces (Giunta et al., 2007). The crosses were characterized by semi-dwarfism and increased earliness, traits that contributed to the replacement of old varieties and landraces. Finally, Green Revolution led to the introgression of *reduced height* (*Rht*) dwarfing genes, enhancing lodging resistance and causing a

* Corresponding author.
E-mail address: masci@unitus.it (S. Masci).

https://doi.org/10.1016/j.fbio.2026.108750
Received 15 January 2026; Received in revised form 23 March 2026; Accepted 26 March 2026
Available online 26 March 2026
2212-4292/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

huge yield improvement (De Vita et al., 2007).

For many years, landraces and historical cultivars have been maintained only by farmers and research institutes, with the risk of contamination or losing them. In recent years, both consumers and the scientific community have shown renewed interest in these ancient genotypes. Modern breeding efforts have substantially increased the productivity and agronomic performance of durum wheat. However, this progress has often come at the expense of genetic diversity, as intensive selection and the repeated use of a narrow genetic base have led to the erosion of valuable alleles present in traditional varieties and landraces (Mazzucotelli et al., 2020; Taranto et al., 2020). Moreover, these varieties represent a useful resource of valuable alleles that need to be exploited yet, especially those associated with resistance to biotic and abiotic stresses (Campanella & Petralia, 2022; Corinzia et al., 2024; Palermo et al., 2024; Ruisi et al., 2015; Scandurra et al., 2024; Scavo et al., 2022).

To promote the preservation of their biodiversity, the Italian government, implementing the European Union directive 2008/62/CE, with the Legislative Decree No. 149 of 2009 (now repealed and replaced by Legislative Decree No. 20 of 2021) permitted the registration of "conservation varieties", defined as "landraces and varieties which are naturally adapted to local and regional conditions and threatened by genetic erosion", in the National catalogue of agricultural species. Since 2009, twenty-eight durum wheat conservation varieties have been inscribed in the Catalogue and twenty-four of them are Sicilian.

Sicily has played a pivotal role in the history of durum wheat cultivation and remains the second-largest Italian region for its production. Here, historical varieties are cultivated in marginal areas under low-input farming systems, partly due to their yield stability (Bosi et al., 2022; Corinzia et al., 2024; De Vita et al., 2010. The products derived from these varieties are positioned in high value and niche markets, where attributes such as unique sensory qualities, nutritional benefits, cultural heritage, and sustainable production practices are highly valued by consumers (Varia et al., 2021).

One of the major challenges in the conservation varieties supply chain is the risk of genetic contamination, which can arise from different factors, such as unintentional seed mixing during the sowing or the harvest, errors in morphological identification or intentional fraud.

One of the main methods for identifying wheat varieties involves analysing the electrophoretic patterns of glutenin subunits and gliadins through SDS-PAGE (Sodium Dodecyl Sulphate-PolyAcrylamide Gel Electrophoresis) and A-PAGE (Acid Poly Acrylamide Gel Electrophoresis) (Bushuk & Zillman, 1978; Lawrence & Shepherd, 1980; Shewry et al., 1978). Visioli et al. (2021) used SDS-PAGE of glutenin subunits to analyse seed batches and commercial flours of Sicilian conservation varieties labelled as mono-varietal and found in all the cases they detected varying degrees of genetic contamination. In a previous work, we demonstrated that gliadin and glutenin electrophoretic patterns can be used to discriminate 21 Sicilian conservation varieties, with the only exception of Bidì, Cappelli and Capeiti 8, that resulted undistinguishable (Palombieri et al., 2024).

Although these techniques are still widely used in wheat research due to their simplicity and cost-effectiveness, the analysis of DNA polymorphisms remains the most effective and high-performance strategy for studying genetic diversity. The earliest wheat genotyping methods were based on Simple Sequence Repeats (SSRs) (Cooke et al., 2003; Pasqualone et al., 2010; Perry, 2004). Advancement in next generation sequencing techniques permitted the discovery of numerous single nucleotide polymorphisms (SNPs) and the development of different DNA microarray platforms (Allen et al., 2017; Burridge et al., 2024; Cavanagh et al., 2013; Rimbert et al., 2018; S. Wang et al., 2014; M. O. Winfield et al., 2016). All these platforms were originally designed on common wheat (*Triticum aestivum* L.) but they have been also used extensively to study durum wheat genetic diversity (Ganugi et al., 2021; Maccaferri et al., 2019; Mazzucotelli et al., 2020; Roncallo et al., 2021; Taranto et al., 2020; Velimirović et al., 2023).

Some authors used these SNP microarrays to study Sicilian tetraploid wheat conservation varieties. Fiore et al. (2019) used the Illumina Wheat 90K microarray to analyse 27 Sicilian tetraploid landraces from the collection of the "Experimental Sicilian Station for Durum Wheat" (ESS) in Caltagirone (Catania, Italy), together with two historical varieties (Cappelli and Trinakria) and two modern varieties. In a recent work Fiore et al. (2022) used the same microarray on a collection that included 104 accessions of *T. turgidum* subspecies, mostly consisting of Sicilian landraces. Also, in this case they were retrieved from ESS collection. Taranto et al. (2022) collected 357 accessions of varieties Timilia reste nere and Russello from Sicilian farmers and preserved them following three generations of Single Seed Descendent (SSD) at the Research Centre for Cereal and Industrial Crops (CREA-CI) of Foggia (Italy). Then, the accessions were genotyped using the Illumina iSelect 15K wheat SNP array (TraitGenetics GmbH, Gatersleben, Germany) to analyse them.

Traceability DNA-based techniques are widely used in food of plant origin supply chains (Lanubile et al., 2024). Over the years, different approaches have been developed for durum wheat varietal identification. SSR markers have been successfully used for samples of seeds, bread and pasta (Fanelli et al., 2023; Pasqualone et al., 2010; Perry, 2004). SNPs have been exploited using a wide range of techniques, as Denaturing High Performance Liquid Chromatography (DHPLC), Cleaved Amplified Polymorphic Sequence (CAPS), TaqMan genotyping, loop-mediated isothermal amplification reaction (LAMP) and Digital PCR (Cibecchini et al., 2020; Giancaspro et al., 2016; Graci et al. 2022; Lee et al., 2023; Morcia et al., 2021).

The aim of this work was to genetically characterize different accessions of six Sicilian tetraploid wheat conservation varieties (Bidì, Ruscìa, Russello, Perciasacchi, Timilia reste bianche and Timilia reste nere) and to develop a DNA-based system to perform traceability through supply chain. These accessions were obtained from custodian farmers, who are actively involved in their conservation and maintenance.

For genetic characterization we performed SNP genotyping using the Wheat 7K Illumina Infinium microarray (Traitgenetics GmbH, Gatersleben, Germany), a platform containing 6707 high polymorphic SNPs selected from previous microarrays.

The resulting data were used to assess the genetic differentiation among the varieties and to identify a subset of SNPs able to discriminate efficiently them.

2. Methods

2.1. Plant materials

A total of 181 accessions of tetraploid wheats were analysed in this work, 141 of which were collected directly from farmers. In each farm, varieties were classified based on the morphology of the spike and two or three spikes for each type were collected. In total, we analysed 20 accessions of Bidì, 31 accessions of Perciasacchi, 15 accessions of Ruscìa, 24 accessions of Russello, 28 accessions of Timilia reste bianche and 23 accessions of Timilia reste nere. Additionally, we included 16 durum wheat modern varieties, 9 durum wheat old cultivars and 15 tetraploid wheat landraces from various sources. The complete accession list is reported in Supplementary Table 1, while Table 1 presents a brief description of the analysed varieties and the corresponding number of accessions.

The collection was composed by 139 accessions of *T. turgidum* spp. *durum* and 42 of *T. turgidum* ssp. *turanicum*. Accessions were divided into landraces (LR), old (OC) and modern cultivars (MC) based on their year of release and their constitution, according to the definition given by Taranto et al. (2020).

Table 1
List of the analysed varieties and their corresponding number of accessions. LR = landrace, OC = old cultivar, MC = modern cultivar.

Variety/cultivar	Breeding type	Subspecies	N° Accessions	Description	Pedigree
Bidì	OC	*durum*	20	Sicilian old cultivar	Selection from the Tunisian landrace Mahmoudi
Anco Marzio	MC	*durum*	-	Modern cultivar	Stot//Altar 84/ALD
Antalis	MC	*durum*	-	Modern cultivar	-
Aureo	MC	*durum*	-	Modern cultivar	Kofa/Svevo
Bronte	MC	*durum*	-	Modern cultivar	Berillo/Latino
Capeiti 8	OC	*durum*	2	Old Italian cultivar released in 1955, registered as conservation variety.	Eiti 6/Cappelli
Cappelli	OC	*durum*	5	Old Italian cultivar released in 1915	Selection from the Tunisian landrace Jenah Rhetifah
Duilio	MC	*durum*	-	Modern cultivar	Cappelli//Anhinga/Flamingo
Emilio Lepido	MC	*durum*	-	Modern cultivar	Orobel//Arcobaleno/Svevo
Etrusco	LR	*turanicum*	5	Turanicum landrace	-
Furio Camillo	MC	*durum*	-	Modern cultivar	Iride/Claudio
Graziella Ra	LR	*turanicum*	1	Turanicum landrace	-
Iride	MC	*durum*	-	Modern cultivar	Altar 84/Ionio
Isola	MC	*durum*	-	Modern cultivar	Rusticano/Platani
K26	MC	*durum*	-	Modern cultivar	Spontaneuos mutant derived from Simeto/Colosseo
QK-77	MC	*turanicum*	-	Turanicum cultivar released in 1990	
Mahmoudi	OC	*durum*	2	Tunisian landrace	-
Marco Aurelio	MC	*durum*	-	Modern cultivar	Orobel//Arcobaleno/Svevo
Opera	MC	*durum*	-	Modern cultivar	Duilio/unspecified CIMMYT line
Orizzonte	MC	*durum*	-	Modern cultivar	Rusticano/Simeto
Perciasacchi	LR	*turanicum*	32	Sicilian turanicum landrace	-
Redidenari	MC	*durum*	-	Modern cultivar	Cofa/N185
Ruscìa	LR	*durum*	15	Sicilian landrace	-
Russello	LR	*durum*	26	Sicilian landrace	-
Samarcanda	MC	*durum*	-	Modern cultivar	Capeiti 8/Rusticano
Saragolla lucana	LR	*turanicum*	3	Turanicum landrace	-
Simeto	MC	*durum*	-	Modern cultivar	Capeiti 8/Valnova
Timilia reste bianche	LR	*durum*	28	Sicilian landrace	-
Timilia reste nere	LR	*durum*	23	Sicilian landrace	-
Trinakria	OC	*durum*	2	Old Italian cultivar released in 1970	B14/Capeiti 8

2.2. DNA extraction and SNP genotyping

One seed from each individual accession was germinated and leaves were sampled for DNA extraction. DNA extraction and SNP genotyping were performed by Traitgenetics GmbH (Gatersleben, Germany). For SNP genotyping, the Illumina Infinium 7K wheat microarray (Traitgenetics GmbH, Gatersleben, Germany) was used, containing 6713 SNP markers.

2.3. Data quality control

Genotyping data were firstly converted into an HapMap file, adding the physical position of each probe on durum wheat genome (Svevo.v1), using R v4.4.3 (R Core Team, 2024). All the probes that did not map were removed. Then, the HapMap file was converted into a Plink file using TASSELL v.5 (Bradbury et al., 2007). Plink v1.9 and Plink v2.0 were used for initial data quality filtering (Chang et al., 2015). Probes mapping on multiple positions were removed using the plink2 function '–rm-dup exclude-all'. Probes mapping on the same physical position were identified (plink –list-duplicate-vars ids-only) and removed. Then, we excluded markers with missing call rates higher than 5% (plink –geno 0.05) and samples with missing call rates higher than 10% (plink –mind 0.1).

2.4. Sample filtering

To identify duplicates accessions and possible misidentifications, the function *mlg.filter* of the R package *poppR* v2.9.6 was used (Kamvar et al., 2015). We used this function to collapse unique multilocus genotypes using a nearest neighbor clustering method calculated on Hamming distance. For this analysis missing data were removed. Accessions of a variety that clustered together with accessions of another variety were removed, because considered misidentifications. For each cluster, only one accession was kept, choosing the one with less missing data. If a cluster contained accessions from multiple varieties, at least one accession from each variety was retained.

To identify more misidentifications, we calculated identity-by-state (IBS) using the plink function '–distance square ibs' and we produced a clustered heatmap representing IBS results using the R package *pheatmap* v1.0.12 (Kolde, 2019). Accessions of a variety that were in a separate cluster than the other accessions of the same variety were removed.

2.5. Population structure and admixture analysis

Markers with a minor allele frequency (MAF) lower than 1% were removed using the plink function '–maf 0.01'. Then, linkage disequilibrium (LD) pruning was made using a r^2 threshold of the 80%, a window size of 50 kb and a step size of 5 (plink –indep-pairwise 50 5 0.8).

The structure of the population and the individual admixture proportions were calculated using the software PopCluster v1.4 (J. Wang, 2024). Best number of population clusters (K value) was determined performing 100 replicate runs (K range from 2 to 20) using two different built-in K estimators: the second order rate of change of the estimated log-likelihood and the F statistics. Then, 1000 replicate runs were made with the ideal number of clusters. Default settings were used except for likelihood scaling, that was set as medium.

The admixture plot was produced in R using the package *ggplot2* v3.5.1 (Wickham, 2016).

Pairwise fixation index (F_{ST}) between populations was calculated using PopCluster and plotted using the R package *corrplot* v0.95 (Wei &

Simko, 2024).

2.6. Diversity analysis

Multidimensional scaling was computed on the pruned SNP dataset using the plink function '–cluster –mds-plot'. The plot was produced using the R package *ggplot2* v3.5.1 (Wickham, 2016).

To study the phylogenetic relationships between the accessions, a neighbor-joining tree on Hamming distance was computed. Hamming distance was calculated using the plink function '–distance square' while the neighbor-joining tree was estimated using the R package ape v5.8 (Paradis & Schliep, 2019). Tree bootstrap values were calculated using poppR v2.9.6 (Kamvar et al., 2015). The tree was plotted using the R package *ggtree* v3.14.0 (Yu et al., 2017).

2.7. SNP statistics

The distribution of markers on the different chromosomes was plotted using the web server SRplot (Tang et al., 2023).

Nei's gene diversity (expected heterozygosis) and the number of private alleles were calculated using the R package *poppR* v2.9.6 (Kamvar et al., 2015). Analysis of molecular variance (AMOVA) between populations was computed using the same package.

2.8. Traceability

A custom R script was developed to identify discriminant markers that are specific to a given variety or population. Briefly, starting from the dataset that had not been filtered for MAF and LD pruned, SNPs with a MAF of 0 within the accessions of the target variety/population were initially selected. This subset of SNPs was then assessed across the entire dataset to identify unique alleles (alleles exclusively present in the variety/population of interest). In cases where no unique alleles were detected, an alternative approach was implemented to identify a set of discriminant probes. First, SNPs that were heterozygous or missing were removed. The MAF was calculated, and only those markers for which the target variety/population carried the minor allele were retained. Finally, SNPs with the lowest MAF values were sequentially selected until a sufficient number was obtained to distinguish the variety/population from all others.

KASP (Kompetitive Allele Specific Polymorphism) markers were designed using Polymarker v1.2.1 (Supplementary Table 2) (Ramirez-Gonzalez et al., 2015). KASP validation was performed on a BioRad CFX duet Real-Time PCR System (Bio-Rad, California, USA). DNA of an accession of each target variety/population was extracted using the method reported in Frittelli et al. (2025). Each KASP reaction was set as follows: 5 μL of KASP-TF V5.0 2X Master Mix (LGC genomics), 0.14 μL of KASP primer mix, 4.86 μL of DNA. KASP primer mix was obtained mixing 30 μL of 0.1 mM common primer with 12 μL of each 0.1 mM allele-specific primer and 46 μL of DNase-free water. Cycling conditions suggested by TAQ manufacturer were used.

2.9. Gliadins extraction and electrophoresis separation

Gliadins were extracted from a bulk of five seeds of one spike of each accession using the method reported in Palombieri et al. (2024). Acid polyacrylamide gel electrophoresis (A-PAGE) was performed using polyacrylamide gels (8 cm × 13 cm) with 8% of T value (total acrylamide concentration) and 1.25% of C value (cross-linker concentration), as reported in Palombieri et al. (2024). As a reference, we used gliadins extracted from the varietal bulks of Sicilian conservation varieties used in Palombieri et al. (2024).

2.10. Statistical analysis

AMOVA significance testing was obtained through a randomization method with 1000 permutations implemented in the *randtest()* function of the R package *ade4* v1.7.23 and a significance level of $\alpha = 0.05$ (Thioulouse et al., 2018).

3. Results

3.1. SNP genotyping

6713 SNPs were characterized in 182 accessions of tetraploid wheat varieties using the Illumina Infinium 7K wheat microarray (Traitgenetics GmbH, Gatersleben, Germany). First step of the analysis was the quality filtering of the obtained markers. A total of 1126 SNPs that did not map on Svevo v.1 genome, map on multiple positions or map on the same position were removed. 444 SNPs were removed because their call rate was below the 95%, leaving a final dataset of 5143 variants. All the samples analysed had a call rate higher of 90%.

To remove duplicates or accessions that were not correctly identified, a multilocus genotype clustering approach was used. Clusters and the list of accessions that were retained or excluded from the analysis are reported in Supplementary Table 3. Clusters 1, 3, 4, 5, 6, 9, 10, 11, 12 and 13 contained multiple accessions of the same variety. In these cases, the choice of accessions to retain was based on the percentage of missing values. In cluster 2 one accession of Timilia reste bianche and one accession of Timilia reste nere were grouped together; given the likely misidentification, we retained only the accession with fewer missing data ("Timiliarb41B″). Several accessions belonging to different *turanicum* varieties (also known as khorasan wheat) (Perciasacchi, Etrusco, Graziella Ra, Saragolla lucana) clustered together in cluster 8; in this case we retained only one accession for each variety. Cluster 7 included four accessions of Bidì, all the accessions of Cappelli, two accessions of Timilia reste bianche and seven accessions of Timilia reste nere. The accessions belonging to "Timilia" varieties were considered misidentifications and removed, while one accession each of Bidì and Cappelli was retained.

A clustered heatmap based on Identity-by-State (IBS) distances among the remaining 126 accessions was generated to identify other potential misidentifications (Supplementary Fig. 1). The accessions "Perciasacchi_13B", "Ruscia_32B″, Ruscia_32C″ and "Timilia_rn_58B″ clustered together with modern and old cultivars and were therefore excluded. Additionally, "Perciasacchi_10B″, "Russello_7" and "Bidì_1C″ clustered separately from all the other accessions and were removed as well.

To validate the putative misidentifications found using SNP analysis, gliadin A-PAGE was performed, and the results are reported in Supplementary Fig. 2. Differences in gliadin composition in comparison with the reference accessions reported in Palombieri et al. (2024) supported the genetic clustering results.

The resulting dataset, consisting of 119 samples, was used for subsequent analyses.

SNP markers were filtered using a minor allele frequency (MAF) threshold of 5%, followed by linkage disequilibrium (LD) pruning with an r^2 threshold of 0.80. The final dataset was composed by 1645 high-quality SNPs distributed across the 14 chromosomes (Fig. 1).

The number of SNPs for each chromosome is reported in Supplementary Table 4. On average, the dataset contained 117 SNPs per chromosome. Chromosome 4A had the lowest number of SNPs (81), while chromosomes 2B and 5B had the highest (145 SNPs).

3.2. Population structure

Population structure was inferred and individual admixture proportions were estimated using an unsupervised likelihood method included in the software PopCluster. The optimal number of population clusters (K = 6) was determined using a maximum likelihood clustering method under a mixture model. Then admixture proportions were calculated using an expectation maximisation algorithm. Every

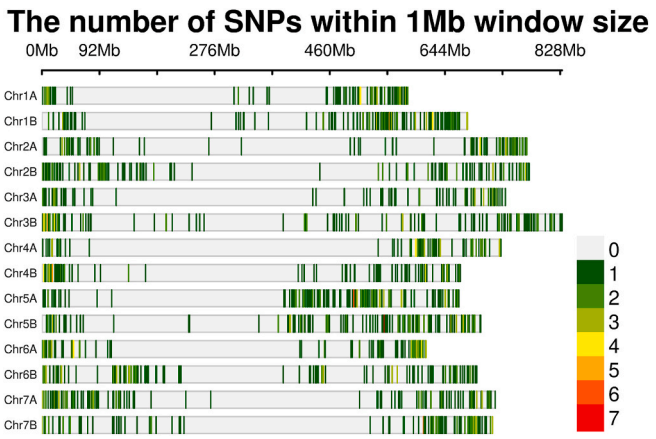


Fig. 1. Distribution of the 1645 high quality SNPs throughout the 14 chromosomes within 1 Mb window size.

accession was considered belonging to a population if the relative admixture proportion was above 0.5. The admixture plot is reported in Fig. 2 and the admixture proportion of each accession are reported in Supplementary Table 5.

The first cluster included all accessions of Timilia reste bianche and Timilia reste nere (n = 36) and was therefore designed as the "Timilia" cluster (Fig. 2, in green). These accessions exhibited minimal with the other populations, with all accessions having an ancestry in this cluster >0.987.

The second cluster included all the Russello accessions (n = 24), while third cluster included all the Ruscìa accessions (n = 11) (Fig. 2, in pink and in purple respectively). Some Russello accessions shared ancestry with the Ruscìa cluster up to 0.269. All Ruscìa accessions shared low ancestry with other populations.

All accessions belonging to *T. turgidum* ssp. *turanicum* (n = 21) clustered together (Fig. 2, in yellow). In this case, the population structure did not turn out to be so distinct, with 4 Perciasacchi accessions showing an ancestry lower than 0.7.

The fifth cluster (n = 12) grouped all the accessions classified as "old cultivars" (OC, defined as the cultivars released before the 1970) along with the modern cultivar, Simeto (released in 1988) (Fig. 2, in orange). Notably, Simeto is the oldest "modern cultivar" included in this study and one of its parental lines is Capeiti 8, an old cultivar. This likely explains its closer genetic affinity to the old cultivars rather than to more recently developed moder varieties.

Finally, the remaining modern cultivars (MC) clustered together (n = 15), forming a distinct group (Fig. 2, in teal green).

To further investigate the degree of genetic differentiation among the six identified populations, pairwise fixation index (F_{ST}) values were calculated (Fig. 3).

The lowest levels of genetic differentiation were observed between OC and MC (F_{ST} = 0.36) and between Ruscìa and Russello (F_{ST} = 0.36). In contrast the highest degree of differentiation was found between the Turanicum and Ruscìa populations (F_{ST} = 0.64). Ruscìa showed a F_{ST}>0.50 with all populations except Russello. Turanicum had highest F_{ST} with Ruscìa, Russello and Timilia than OC and MC. MC and Russello showed tendentially lower fixation indexes.

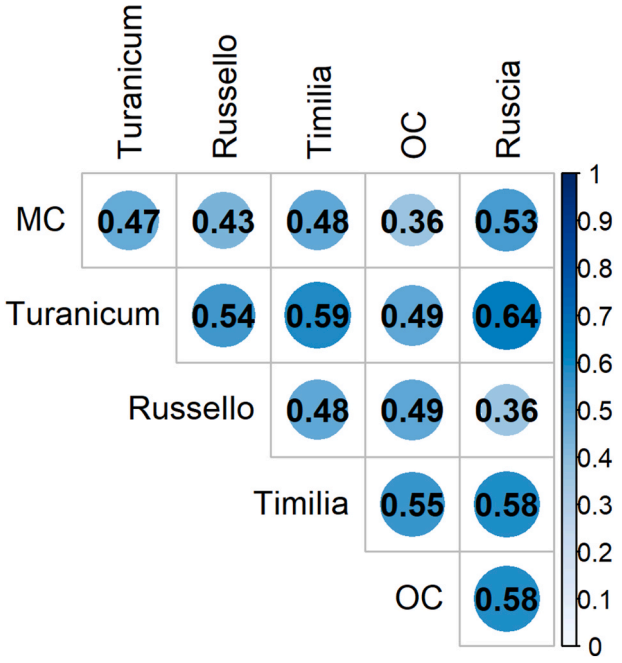


Fig. 3. Pairwise fixation index (Fst) between the six population. OC = old cultivars, MC = modern cultivars.



Fig. 2. Admixture plot (K = 6) of 119 accessions analysed using 1645 markers. Each vertical bar represents one accession. In each bar, the colour represents the ancestry (admixture proportion) relative to one of the six populations. OC = old cultivars, MC = modern cultivars.

3.3. Diversity analysis

A multidimensional scaling (MDS) analysis was performed on the dataset composed by 119 accessions and 1645 markers to visualize the genetic relationships and structure within the population (Fig. 4).

The MDS analysis confirmed that the "Timilia" accessions formed a distinct cluster completely separated from all other populations, with no distinction between Timilia reste bianche and Timilia reste nere. The Russello and Ruscìa populations were overlapped but remained distinct from the other populations. The clusters representing OC and MC were almost completely overlapped, probably due to the share of the same genetic background. Surprisingly, the Turanicum cluster partially overlapped with OC and MC. Phylogenetic relationships among the accessions were furtherly studied constructing a neighbor-joining tree on Hamming distance. For improved readability, Fig. 5 presents the tree without boostrap values, while Supplementary Fig. 3 shows the same tree with bootstrap values on the nodes.

Phylogenetic tree confirmed the genetic distance of Timilia accessions from all the other accessions included in the study. No clear clustering between Timilia reste bianche and Timilia reste nere was found.

Ruscìa and Russello accessions share a common ancestral node, but they form two distinct clades.

All *T. turgidum* ssp. *turanicum* clustered together but no clear distinction among the accessions was found. In fact, Perciasacchi accessions did not cluster all together, indicating intra-variety variability. Notably some Etrusco, Saragolla Lucana, Graziella Ra and Perciasacchi accessions resulted identical (Supplementary Table 1).

The MC and OC populations shared a common ancestral node in the phylogenetic tree. Within the OC population, two distinct sub-clades were identified. The first one included Bidì, Cappelli and Mahmoudi accessions, historical varieties derived from Tunisian landraces. The second one encompassed Capeiti 8 and Trinakria (which are descendent of Cappelli) and Simeto (derived from a cross involving Capeiti 8). It is also noteworthy that some Cappelli and Bidì accessions were identical (Supplementary Table 1).

3.4. Statistics

To further study populations diversity, Nei's gene diversity (H_{exp}), Fixation index (F_{ST}) and the number of private alleles were calculated (Table 2).

Nei's gene diversity (also called expected heterozygosity) measure the average genetic diversity per locus and is related to intra-population variability. Nei's gene diversity was higher in MC (H_{exp} = 0.235) and Russello (H_{exp} = 0.201), followed by OC (H_{exp} = 0.152), Timilia (H_{exp} = 0.139), Turanicum (H_{exp} = 0.126) and Ruscìa (H_{exp} = 0.112) populations.

Fixation index (F_{ST}) gives information about the degree of differentiation of the populations. F_{ST} had an opposite trend in comparison with Nei's gene diversity, with values ranging from 0.201 to 0.614. MC had the major number of private alleles (n = 80), followed by Russello (n = 45) and Timilia (n = 42). The lowest number of private alleles were found in Turanicum (n = 17), OC (n = 11) and Ruscìa (n = 5).

An analysis of Molecular Variance (AMOVA) was conducted to assess the genetic structure of the populations and the results are reported in Table 3.

AMOVA revealed that most of the variation in the analysed dataset was due to population structure accounting for 50.82%) of the total variance (p-value<0.001). A similar proportion of variation (46.94%, p-value <0.001) was explained by differences between samples within populations. The remaining 2.24% of variation was attributed to sample variation, that represents the error of the model.

3.5. Detection of markers with potential traceability applications

Using a custom R script, a total of 18 markers were found able to discriminate the varieties Bidì (n = 5), Ruscìa (n = 2) and Russello (n = 2) and the two "populations" Timilia (n = 2) and Turanicum (n = 7) (Supplementary Table 2). Unique alleles were found only for Timilia, while for the other varieties/populations, discrimination could be achieved with the combination of markers. However, we were unable to find markers that could distinguish between Timilia reste bianche and Timilia reste nere, so these two varieties were treated as a single group. The same issue occurred with the Turanicum accessions. Additionally, some Bidì accessions were identical to Cappelli, making it impossible to discriminate between these two varieties.

After identifying the 18 markers, we proceeded to design KASP (Kompetitive Allele-Specific PCR) assays to enable bi-allelic scoring, with the goal of achieving a rapid and cost-effective genotyping method. The assays were designed using the Polymarker software, which facilitates the development of homeoallele-specific markers (Supplementary Table 2). Their efficiency and specificity were empirically validated using a Bio-Rad CFX Duet Real-Time PCR system. All assays demonstrated clear allele discrimination among the tested samples.

4. Discussions

After having been abandoned for years, the cultivation of wheat conservation varieties is spreading again, due mainly to the interest of the consumers. Durum wheat is the most cultivated cereal in Italy, representing one of the most important crops economically and

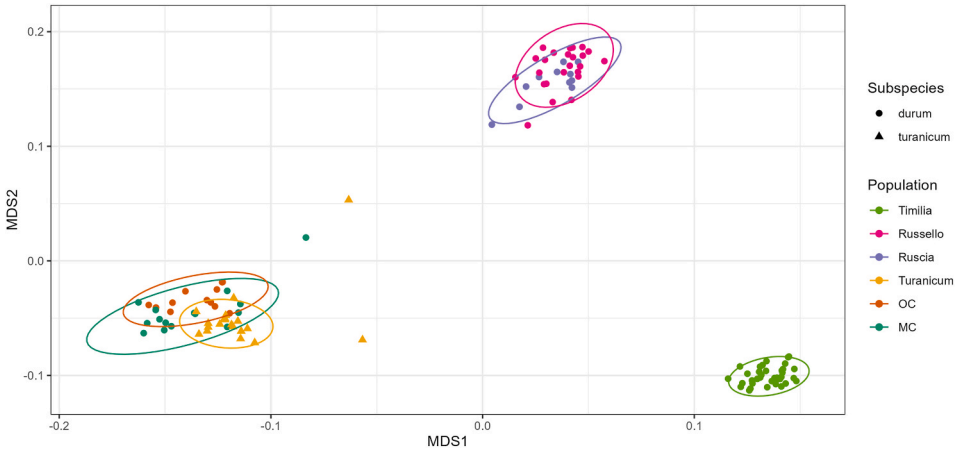


Fig. 4. Multidimensional scaling (MDS) plot of the 119 accessions with 1645 markers. Populations identified by PopCluster are marked with different colours. Different point shape indicates the subspecies. Ellipses represent the 95% confidence interval of each population.

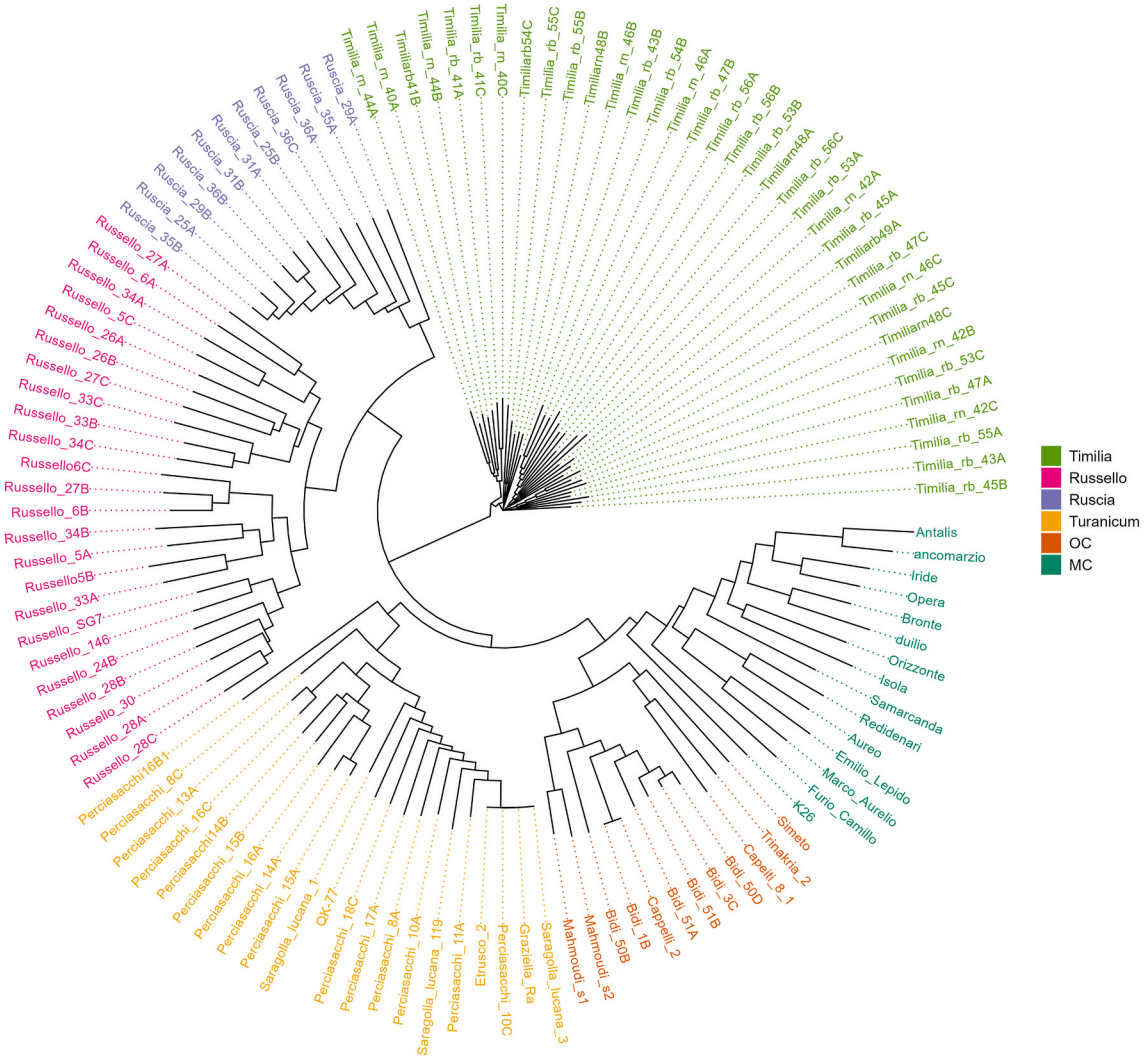


Fig. 5. Neighbor-joining tree based on Hamming distance of 119 accessions analysed using 1645 markers. Different colours represent different populations.

Table 2
Statistics of the different populations. N = number of individuals, H_{exp} = Nei's gene diversity (expected heterozygosity), F_{ST} = Fixation index. Private alleles are the alleles that occurs only in one population.

Population	N	H_{exp}	F_{ST}	Private alleles
MC	15	0.235	0.201	80
OC	12	0.152	0.476	11
Ruscìa	11	0.112	0.614	5
Russello	24	0.201	0.319	45
Timilia	36	0.139	0.535	42
Turanicum	21	0.126	0.576	17

Table 3
AMOVA results using 1645 markers of 119 accessions and considering 6 populations. Df = degree of freedom.

Source of variation	Df	Sum of squares	Percentage of variation	p-value
Between populations	5	55737.77	50.815	<0.001
Between samples within populations	113	59470.39	46.943	<0.001
Within samples	119	1461.07	2.243	
Total	237	116669.24	100	

traditionally. Its history is tightly linked to this country, from the cultivation of landraces to the first breeding programs at the beginning of the 20th century. So, it is unsurprising that Italy has a richness in tetraploid wheat conservation varieties. They are particularly important in Sicily, where they are appreciated for their ability to be productive in marginal areas.

For years, these varieties have been preserved in research institutes collections, but they have been mainly maintained and disseminated by farmers. This often led to varietal contamination, due to accidental seeds mixing pre- or post-harvest. Visioli et al. (2021), analysing glutenin composition, found a degree of contamination ranging from 26% to 53% in conservation varieties grain batches provided by Sicilian farmers.

SNP microarray is a cost-effective and high-throughput technique to study genetic diversity in wheat. In the last year, many of them were developed and used for applications as genomic-wide association studies (GWAS), quantitative trait loci (QTL) mapping, marker assisted selection (MAS) or population genetics studies (Esposito et al., 2024; Liu et al., 2024; Mazzucotelli et al., 2020; Sertse et al., 2023; P. Wang et al., 2022).

Several authors used SNP microarray to characterize wheat landraces and old cultivars (Ganugi et al., 2021; Mangini et al., 2018; Miazzi et al., 2022; Negisho et al., 2021; Taranto et al., 2020; Velimirović et al., 2023). Given the importance of wheat conservation varieties in Sicily, this technique was used to characterize them in several works (Fiore et al., 2019, 2022; Taranto et al., 2022).

In our work, we collected numerous accessions of Sicilian tetraploid

wheat conservation varieties directly from farmers responsible for their maintenance and commercialization with the aim to genetically characterize them using the Illumina Infinium 7K wheat microarray (Traitgenetics GmbH, Gatersleben, Germany). To capture as much of the heterogeneity of these varieties as possible, we randomly collected spikes representing the different morphological types identified within each farmer's plot. A sampling strategy was necessary, given the impossibility to analyse all the plants; however, this might represent a limitation of this study, as some genotypes may not be included. Nevertheless, sampling multiple morphological type across different farmers allowed us to obtain a representative sample of each variety. Together with Sicilian conservation varieties, we also collected other tetraploid wheat landraces, historical and modern varieties to use as a comparison. The final aim was to assess the diversity present today in these varieties and to develop a traceability method, useful for the valorisation of the supply chain.

The initial sample set was composed by 181 accessions, 141 of which were Sicilian conservation varieties. The first step of the analysis was the identification of duplicate or incorrectly identified samples. As stated above, it is common to find contaminated batch of seeds of wheat conservation varieties. Already during sample collection in the field, we observed variation in spikes morphology, attributable either to natural variability, particularly in the case of landraces, or to possible contamination. In the first case, spikes of each morphology type were collected; in the second case spikes were not collected. We used a multilocus genotype clustering and an identity-by-state (IBS) approach to identify and remove 45 duplicate and 17 misidentified accessions. Both approaches were based on Hamming distance calculation: the first was to find genetically identical individuals also in presence of missing data, the second was used to find accessions of a variety more similar to accessions of other varieties. Several instances of misidentification were identified during the analysis; in one case one accession of Timilia reste bianche was found to be identical to another of Timilia reste nere. These two varieties derive from the same population "Timilia" and are mainly distinguished by the awn colour, yellowish in the first and black in the second (De Cillis, 1942). They are often confused, especially since awn colour can be influenced by environmental factors. In the second case, seven accessions of Timilia reste nere and two of Timilia reste bianche were removed because of their identity with Bidì and Cappelli accessions. Another Timilia reste nere accession (Timilia_rn_58B) was discarded because it clustered with Bidì and Cappelli accessions. Interestingly, these 10 accessions were sampled from two fields of two different farmers (Supplementary Table 1), highlighting the presence of contamination in different batches. In the third case, one Russello, two Ruscìa, one Bidì and three Perciasacchi accessions were removed because falling outside the cluster of the relative varieties. Gliadin A-PAGE was performed to verify the suspected misidentifications and to avoid discarding potential sources of variability.

Most of the duplicated accessions identified in the analysis belonged to the same variety, with two exceptions. In one case, all Cappelli accessions were not only identical to each other but also matched four Bidì accessions.

Bidì and Cappelli are morphologically very similar and are often confused with one another. The four Bidì accessions (Bidi_1A, Bidi_1B, Bidi_52A, Bidi_52B) were collected from two different farmers but exhibited an elongated spike, which is atypical for Bidì. This suggests that they have been misidentified. However, to retain valuable information, we decided to keep one of these accessions (Bidi_1B) based on its minimal missing data. In the second case, we found that fourteen Perciasacchi, four Etrusco, one Graziella Rà and one Saragolla Lucana accessions were genetically identical. We decided to keep only one accessions for each variety, prioritizing those with the least missing data.

The final dataset composed of 119 accessions was used for genetic diversity analysis. The first step was the inferring of the population structure using a maximum likelihood clustering approach. An ideal number of six clusters was identified that represented the following populations: "Timilia", "Russello", "Ruscìa", "Turanicum", "Old cultivars" and "Modern cultivars". The inferred populations were used for the subsequent analyses. AMOVA analysis validated this population structure. In fact, 50.82% of the variance found in our dataset was given by difference between populations, while 46.94% of the variance was given by differences between samples within populations, denoting also intra-population high variability. High pairwise F_{ST} index between populations (>0.36) further confirmed the calculated population structure.

The "Timilia" population contained 23 accessions of Timilia reste bianche and 13 accessions of Timilia reste nere. Neither multidimensional scaling plot nor neighbor-joining tree were able to discriminate between these two varieties. A relatively low Nei's gene diversity index was found ($H_{exp} = 0.139$), indicating low intra-population variability. This value was lower than the one ($H_{exp} = 0.167$) found by Taranto et al. (2022) for the Timilia collection (n = 157). This discrepancy is probably due to the lower number of individuals present in our study. As stated before, these two varieties were selected from the same population based on the awn colour. The results of our study confirm a strong genetic correlation between these two varieties. The impossibility to discriminate them could be attributed to the absence of appropriate markers in our dataset. Palombieri et al. (2024) found differences in ω-gliadins and low molecular weight glutenin subunits (LMW-GS) between the two varieties, suggesting genetic variation in the group 1 chromosomes. Interestingly, genes controlling awn colour are located in the same chromosomes group (Gadaleta et al., 2009). Admixture analysis indicated a very low degree of genetic mixings with other populations. Furthermore, the multidimensional scaling plot and the neighbor-joining tree highlighted a high genetic distance from other populations. Taranto et al. (2020) analysed 259 tetraploid wheat accessions using the Illumina iSelect 15K wheat SNP array, finding high distance between the nine Timilia accessions included in that study and the rest of the germplasm. Similar results were obtained by Fiore et al. (2022), that analysed 126 tetraploid wheats, including 14 Timilia accessions. All these findings collectively confirm the unique genetic origin of this landrace, suggesting that it has likely never been incorporated in durum wheat breeding programs. Given its distinct genetic profile and the potential insights it could offer, future studies should investigate its applicability and potential contributions to durum wheat breeding, especially in areas such as stress resistance, yield improvement, or quality traits.

Russello and Ruscìa accessions formed two separate populations, with some degree of admixture between each other. Ruscìa, also named "Russello ibleo", is very similar to Russello, often leading to confusion between them. Neighbor-joining tree showed that these two varieties derived from the same ancestral node. Moreover, multidimensional scaling analysis revealed an almost complete overlapping of the two varieties clusters. Pairwise F_{ST} index between the two populations was equal to 0.36, one of the lowest found in this study. Russello population was characterized by a higher Nei's gene diversity index (0.201 vs 0.112 of Ruscìa), a lower F_{ST} index (0.319 vs 0.614 of Ruscìa) and a higher number of private alleles (45 vs 5 of Ruscìa). These data highlight that Russello is characterized by higher intra-population variability than Ruscìa.

All accessions of *T. turgidum* ssp. *turanicum* clustered together, highlighting their genetic similarity. Interestingly, all Perciasacchi accessions did not cluster separately from the other *turanicum* varieties. In the neighbor-joining tree two different clusters emerged: one containing only Perciasacchi accessions and one containing some Perciasacchi accessions alongside those from the other *turanicum* wheat. Additionally, two Perciasacchi accessions exhibited unique clustering behaviour, as evidenced by their separation in both the neighbor-joining tree and the multidimensional scaling plot. This suggests potential genetic diversity within the Perciasacchi accessions, warranting further investigation into their genetic relationships and distinctions within the *turanicum*

subspecies. In this, Perciasacchi cluster overlaps with OC and MC clusters. In our work, the Turanicum accessions showed more similarity with durum wheat old cultivar and modern cultivar than with Sicilian landraces. This result is in accordance with Fiore et al. (2022), where four Turanicum accessions, including two Perciasacchi (called by the synonym Farro lungo), shared the same ancestral node with old and modern cultivars. As said above, we found identity between fourteen Perciasacchi, four Etrusco, one Graziella Ra and one Saragolla Lucana accessions. All these Turanicum varieties originate from Italy, probably sharing the same genetic background. At the moment, we cannot determine whether the elevated identity among these varieties is due to actual homonymy, misidentification, or the lack of adequate markers for discriminating the Turanicum varieties. It is important to consider that wheat microarray platforms and SNP dataset, such as those used in this study, were developed based mainly on *T. aestivum* and only marginally on *T. turgidum* ssp. *durum* (S. Wang et al., 2014; M. O. Winfield et al., 2016). In the case of *T. turgidum* ssp. *turanicum*, our results may be influenced by the limited resolution of the SNP array used and should be interpretated cautiously. Future studies should focus on Next Generation Sequencing (NGS) approaches, which could better elucidate the genetic diversity characterizing these varieties.

OC cluster was composed by Bidì, Capeiti 8, Cappelli, Mahmoudi, Trinakria and Simeto accessions. Mahmoudi is a Tunisian landrace and Bidì and Cappelli are old cultivars selected at the beginning of the 1900 from Tunisian landraces. Neighbor-joining tree showed that they share the same ancestral node. It is commonly that Cappelli was selected by a Tunisian landrace called Jenah Retifah, while Bidì was selected from Mahmoudi. However, De Cillis (1927) reported that Bidì and Cappelli were selected from the same Tunisian landrace. Recent studies seem to support this hypothesis and also our data shows the genetic similarity between Mahmoudi, Bidì and Cappelli (Fiore et al., 2022; Marzario et al., 2018; Taranto et al., 2020). Capeiti 8, Trinakria and Simeto are cultivars developed within Italian breeding programs and share a common genetic origin through the historical variety Cappelli. In fact, Capeiti 8, released int 1955, originates from a cross between Cappelli and Eiti 6. Trinakria, released in 1970, derives from a cross between Capeiti 8 and B14. Finally, Simeto derives from a cross between Capeiti 8 and Valnova and was released in 1988. These genealogical relationships explain the genetic similarity observed among these cultivars. Consistently, the "old cultivars" (OC) and "modern cultivars" (MC) clusters showed a low level of genetic differentiation, as indicated by a pairwise F_{ST} value of 0.36, the lowest observed among the analysed groups. Furthermore, these two clusters largely overlapped in the multidimensional scaling (MDS) analysis, further supporting their shared genetic background.

Last cluster (MC) contained all the moder cultivars included in this study. These are all cultivars released in the 21st century and they are still cultivated in Italy. This cluster showed the highest Nei's gene diversity index (H_{exp} = 0.235) and number of private alleles (n = 80), probably due to its high heterogeneity. The fixation index was the lowest (F_{ST} = 0.201), indicating less differentiation compared to the other populations.

Summing up, we analysed the diversity of Sicilian tetraploid wheat conservation varieties, finding high inter- and intra-populations variability. Notably, high variability and genetic distance were found between Sicilian landraces and modern cultivars, especially in the case of Timilia, highlighting the opportunity to include them in future breeding programs. Our findings complement and expand upon the previous study by Taranto et al. (2022) and Fiore et al. (2022), providing deeper insights into the phylogenetic relationships among Sicilian conservation varieties and reinforcing their relevance for conservation and pre-breeding efforts. Moreover, we investigated the genetic relationship between Italian Turanicum accessions, obtaining unexpected results. Specifically, we found genetic complete identity among accessions of Etrusco, Graziella Ra, some Saragolla Lucana and some Perciasacchi samples. These varieties are also morphologically indistinguishable, which complicates the assessment of whether the observed identity is due to actual synonymy, misidentification, or a lack of discriminating markers. The cultivar QK-77 also showed high genetic similarity to other Turanicum accessions studied. Overall, our findings suggest a shared genetic origin among these varieties and highlight the need for further studies to investigate the global genetic relationships within the *turanicum* subspecies.

Another objective of our work was the development of a DNA-based traceability system for Sicilian tetraploid wheat conservation varieties. KASP assay is a simple and convenient way to perform allelic discrimination, that can be applied to all plant species (Carrara et al., 2023; Sales et al., 2023; Tang et al., 2022; Winfield et al., 2020). We decided to select the minimum number of SNPs that were able to distinguish among the varieties Bidì, Russello and Ruscìa and the accessions belonging to Timilia population and *turanicum* subspecies. The high level of intra-varietal genetic variability observed in conservation varieties represents a major challenge for molecular traceability. To overcome this limitation, multiple accessions per variety were collected from different farmers, and SNPs monomorphic within each population, but polymorphic across populations, were selected. We found a total of 18 SNP markers suitable for this purpose, which were subsequently converted into KASP assays. As highlighted by our findings, misidentification is recurring issue within the conservation varieties supply chain.

From a consumer standpoint, ensuring the authenticity of raw materials is essential, especially when products command a premium price due to their traditional or geographical origin. The set of markers we developed can be applied across different points of the supply chain to prevent fraud, support traceability, and enhance product value. KASP markers offers a faster and more cost-effective technique for varietal identification compared with the currently available methods, such as gliadin A-PAGE and SSR markers. The markers we developed can already be used to genotype seed lots before sowing or milling to guarantee authenticity. Future works will focus on validating the method directly on food products.

Sicilian farmers have an important role for the preservation and valorisation of conservation varieties. Our findings highlight the great genetic diversity they have maintained, but also showed a certain degree of genetic contamination. The development of a DNA-based varietal identification method could support farmers in valorising and improving the purity of their production, as well as protecting it against potential fraud.

5. Conclusion

The recent re-discovery of wheat conservation varieties has led to an increase in their cultivation, especially in Sicily, and represent a valuable opportunity to boost the local agricultural economy. With the aim to study the genetic diversity of these varieties and to identify a method to ensure their authenticity, we performed the genotyping of over one hundred accessions of the most important Sicilian conservation varieties, directly collected from farmers.

Our analyses revealed the presence of unexploited genetic variability in Timilia, Russello and Ruscia accessions. In several cases, genetic contamination was detected, highlighting the need for the development of a traceability method. For this reason, we developed a DNA-based method, based on KASP assays, able to detect the presence of varieties Bidì, Russello, Ruscìa and Timilia.

The findings of our work highlight the necessity of use genotyping to explore the genetic diversity of conservation varieties and provide additional tools to valorise their supply chain, while ensuring the authenticity of their products for consumers.

CRediT authorship contribution statement

Marco Bonarrigo: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal

analysis, Data curation, Conceptualization. **Giulio Metelli:** Writing – review & editing, Methodology. **Bernardo Messina:** Writing – review & editing, Resources, Project administration. **Giuseppe Russo:** Writing – review & editing, Resources, Project administration. **Francesco Sestili:** Writing – review & editing. **Stefania Masci:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization.

Funding sources

This work is part of CIRCE project, funded by PSR SICILIA MISURA 19.1/16.1 of GAL ISC MADONIE (CUP F72C21000270009) and belong to the EIP AGRI (Agricultural European Innovation Partnership), as a part of EU Common Agricultural Policy (CAP) and works to foster competitive and sustainable farming and forestry that "achieves more and better from less".

This research was also carried out within the Agritech National Research Center and received funding from the European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). ENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

The research was carried out within the framework of the Ministry for Education, University, and Research (MIUR) initiative "Department of Excellence" (Law 232/2016) DAFNE Project 2023–27 ″Digital, Intelligent, Green and Sustainable (acronym: D.I.Ver.So.)"

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thanks Maurizio Nigliaccio (Consorzio di Ricerca Gian Pietro Ballatore) for its valuable assistance and Oriana Porfiri (CGS Sementi) and Ines Yacoubi (Centre of Biotechnology of Sfax) for providing wheat accessions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fbio.2026.108750.

Data availability

Data will be made available on request.

References

Allen, A. M., Winfield, M. O., Burridge, A. J., Downie, R. C., Benbow, H. R., Barker, G. L. A., Wilkinson, P. A., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., Pirani, A., Webster, T., Brew, F., Bloor, C., Griffiths, S., Bentley, A. R., Alda, M., Jack, P., … Edwards, K. J. (2017). Characterization of a wheat breeders' array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnology Journal, 15*(3), 390–401. https://doi.org/10.1111/pbi.12635

Blanco, A. (2024). Structure and trends of worldwide research on durum wheat by bibliographic mapping. *International Journal of Plant Biology, 15*(1), 132–160. https://doi.org/10.3390/ijpb15010012

Bosi, S., Negri, L., Fakaros, A., Oliveti, G., Whittaker, A., & Dinelli, G. (2022). GGE biplot analysis to explore the adaption potential of Italian common wheat genotypes. *Sustainability, 14*(2), 897. https://doi.org/10.3390/su14020897

Bożek, K. S., Żuk-Gołaszewska, K., Bochenek, A., Gołaszewski, J., & Kalaji, H. M. (2021). Modelling the growth, development and yield of *Triticum durum* desf under the changes of climatic conditions in north-eastern Europe. *Scientific Reports, 11*(1), Article 21753. https://doi.org/10.1038/s41598-021-01273-8

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Burridge, A. J., Winfield, M., Przewieslik-Allen, A., Edwards, K. J., Siddique, I., Barral-Arca, R., Griffiths, S., Cheng, S., Huang, Z., Feng, C., Dreisigacker, S., Bentley, A. R., Brown-Guedira, G., & Barker, G. L. (2024). Development of a next generation SNP genotyping array for wheat. *Plant Biotechnology Journal, 22*(8), 2235–2247. https://doi.org/10.1111/pbi.14341

Bushuk, W., & Zillman, R. R. (1978). Wheat cultivar identification by gliadin electrophoregrams. I. apparatus, method and nomenclature. *Canadian Journal of Plant Science, 58*(2), 505–515. https://doi.org/10.4141/cjps78-076

Campanella, V., & Petralia, R. (2022). Resistance assessment of durum wheat landraces to fusarium foot rot. *Physiological and Molecular Plant Pathology, 121*, Article 101879. https://doi.org/10.1016/j.pmpp.2022.101879

Carrara, I., Terzi, V., Ghizzoni, R., Delbono, S., Tumino, G., Crespan, M., Gardiman, M., Francia, E., & Morcia, C. (2023). A molecular toolbox to identify and quantify grape varieties: On the trace of "Glera.". *Foods, 12*(16). https://doi.org/10.3390/foods12163091. Article 16.

Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G. L., Akhunova, A., See, D., Bai, G., Pumphrey, M., Tomar, L., Wong, D., Kong, S., Reynolds, M., Da Silva, M. L., Bockelman, H., … Akhunov, E. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proceedings of the National Academy of Sciences, 110*(20), 8057–8062. https://doi.org/10.1073/pnas.1217133110

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience, 4*(1). https://doi.org/10.1186/s13742-015-0047-8. s13742-015-0047–0048.

Cibecchini, G., Cecere, P., Tumino, G., Morcia, C., Ghizzoni, R., Carnevali, P., Terzi, V., & Pompa, P. P. (2020). A fast, naked-eye assay for varietal traceability in the durum wheat production chain. *Foods, 9*(11), 1691. https://doi.org/10.3390/foods9111691

Cooke, R. J., Bredemeijer, G. M. M., Ganal, M. W., Peeters, R., Isaac, P., Rendell, S., Jackson, J., Röder, M. S., Korzun, V., Wendehake, K., Areshchenkova, T., Dijcks, M., Laborie, D., Bertrand, L., & Vosman, B. (2003). Assessment of the uniformity of wheat and tomato varieties at DNA microsatellite loci. *Euphytica, 132*(3), 331–341. https://doi.org/10.1023/A:1025046919570

Corinzia, S. A., Caruso, P., Scandurra, A., Anastasi, U., Cosentino, S. L., & Testa, G. (2024). Yield response and leaf gas exchange of Sicilian wheat landraces. *Agronomy, 14*(5), 1038. https://doi.org/10.3390/agronomy14051038

De Cillis, E. (1927). *I Grani d'Italia. Tipografia della Camera dei deputati*.

De Cillis, U. (1942). I Frumenti siciliani. *Giuseppe maimone editore*.

De Vita, P., Mastrangelo, A. M., Matteu, L., Mazzucotelli, E., Virzì, N., Palumbo, M., Storto, M. L., Rizza, F., & Cattivelli, L. (2010). Genetic improvement effects on yield stability in durum wheat genotypes grown in Italy. *Field Crops Research, 119*(1), 68–77. https://doi.org/10.1016/j.fcr.2010.06.016

De Vita, P., Nicosia, O. L. D., Nigro, F., Platani, C., Riefolo, C., Di Fonzo, N., & Cattivelli, L. (2007). Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *European Journal of Agronomy, 26*(1), 39–53. https://doi.org/10.1016/j.eja.2006.08.009

Esposito, S., Palombieri, S., Vitale, P., Angione, G., D'Attilia, C., Taranto, F., Sestili, F., & De Vita, P. (2024). Identification and development of functional markers for purple grain genes in durum wheat (*Triticum durum* desf.). *Theoretical and Applied Genetics, 137*(9), 210. https://doi.org/10.1007/s00122-024-04710-0

Fanelli, V., Dellino, M., Taranto, F., De Giovanni, C., Sabetta, W., De Vita, P., & Montemurro, C. (2023). Varietal identification in pasta through an SSR-based approach: A case study. *Journal of the Science of Food and Agriculture, 103*(11), 5521–5528. https://doi.org/10.1002/jsfa.12631

Fiore, M. C., Blangiforti, S., Preiti, G., Spina, A., Bosi, S., Marotti, I., Mauceri, A., Puccio, G., Sunseri, F., & Mercati, F. (2022). Elucidating the genetic relationships on the original old Sicilian *triticum* spp. Collection by SNP genotyping. *International Journal of Molecular Sciences, 23*(21), Article 13378. https://doi.org/10.3390/ijms232113378

Fiore, M., Spina, B., Venora, Dell'Acqua, Lupini, P., Monti, P., & Sunseri. (2019). High-throughput genotype, morphology, and quality traits evaluation for the assessment of genetic diversity of wheat landraces from sicily. *Plants, 8*(5), 116. https://doi.org/10.3390/plants8050116

Frittelli, A., Botticella, E., Palombieri, S., Metelli, G., Masci, S., Silvestri, M., Lafiandra, D., & Sestili, F. (2025). Improving the agronomic performance of high-amylose durum wheat. *Plant Science, 355*, Article 112459. https://doi.org/10.1016/j.plantsci.2025.112459

Gadaleta, A., Giancaspro, A., Giove, S. L., Zacheo, S., Mangini, G., Simeone, R., Signorile, A., & Blanco, A. (2009). Genetic and physical mapping of new EST-Derived SSRs on the A and B genome chromosomes of wheat. *Theoretical and Applied Genetics, 118*(5), 1015–1025. https://doi.org/10.1007/s00122-008-0958-1

Ganugi, P., Palchetti, E., Gori, M., Calamai, A., Burridge, A., Biricolti, S., Benedettelli, S., & Masoni, A. (2021). Molecular diversity within a mediterranean and European panel of tetraploid wheat (*T. Turgidum* subsp.) landraces and modern germplasm inferred using a high-density SNP array. *Agronomy, 11*(3), 414. https://doi.org/10.3390/agronomy11030414

Giancaspro, A., Colasuonno, P., Zito, D., Blanco, A., Pasqualone, A., & Gadaleta, A. (2016). Varietal traceability of bread 'Pane Nero di Castelvetrano' by denaturing high pressure liquid chromatography analysis of single nucleotide polymorphisms. *Food Control, 59*, 809–817. https://doi.org/10.1016/j.foodcont.2015.07.006

Giunta, F., Motzo, R., & Pruneddu, G. (2007). Trends since 1900 in the yield potential of italian-bred durum wheat cultivars. *European Journal of Agronomy, 27*(1), 12–24. https://doi.org/10.1016/j.eja.2007.01.009

Graci, S., Olivieri, F., & Barone, A. (2022). Phenotypic and genotypic characterization to authenticate and trace a high-quality yellow tomato ecotype through the processing chain. *Scientia Horticulturae, 306*, Article 111449. https://doi.org/10.1016/j.scienta.2022.111449

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics, 6*. https://www.frontiersin.org/articles/10.3389/fgene.2015.00208.

Kolde, R. (2019). *Pheatmap: Pretty heatmaps*. https://CRAN.R-project.org/package=pheatmap.

Lanubile, A., Stagnati, L., Marocco, A., & Busconi, M. (2024). DNA-based techniques to check quality and authenticity of food, feed and medicinal products of plant origin: A review. *Trends in Food Science & Technology, 149*, Article 104568. https://doi.org/10.1016/j.tifs.2024.104568

Lawrence, G., & Shepherd, K. (1980). Variation in glutenin protein subunits of wheat. *Australian Journal of Biological Sciences, 33*(2), 221. https://doi.org/10.1071/BI9800221

Lee, S.-J., Demeke, T., Dusabenyagasani, M., Saydak, D., Perry, D., & Walkowiak, S. (2023). Evaluation of two high-throughput genotyping systems for rapid identification of Canadian wheat varieties. *Canadian Journal of Plant Science, 103*(4), 422–425. https://doi.org/10.1139/cjps-2022-0192

Liu, L., Qu, P., Zhou, Y., Li, H., Liu, Y., Zhang, M., Zhang, L., Zhao, C., Zhang, S., & Pang, B. (2024). Consensus linkage map construction and QTL mapping for eight yield-related traits in wheat using the BAAFS wheat 90K SNP array. *Journal of Integrative Agriculture, 23*(11), 3641–3656. https://doi.org/10.1016/j.jia.2023.07.028

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., Ormanbekova, D., Lux, T., Prade, V. M., Milner, S. G., Himmelbach, A., Mascher, M., Bagnaresi, P., Faccioli, P., Cozzi, P., Lauria, M., Lazzari, B., Stella, A., Manconi, A., … Cattivelli, L. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics, 51*(5), 885–895. https://doi.org/10.1038/s41588-019-0381-3

Mangini, G., Nigro, D., Margiotta, B., De Vita, P., Gadaleta, A., Simeone, R., & Blanco, A. (2018). Exploring SNP diversity in wheat landraces germplasm and setting of a molecular barcode for fingerprinting. *Cereal Research Communications, 46*(3), 377–387. https://doi.org/10.1556/0806.46.2018.033

Marzario, S., Logozzo, G., David, J. L., Zeuli, P. S., & Gioia, T. (2018). Molecular genotyping (SSR) and agronomic phenotyping for utilization of durum wheat (*Triticum durum* Desf.) Ex situ collection from Southern Italy: A combined approach including pedigreed varieties. *Genes, 9*(10). https://doi.org/10.3390/genes9100465. Article 10.

Mazzucotelli, E., Sciara, G., Mastrangelo, A. M., Desiderio, F., Xu, S. S., Faris, J., Hayden, M. J., Tricker, P. J., Ozkan, H., Echenique, V., Steffenson, B. J., Knox, R., Niane, A. A., Udupa, S. M., Longin, F. C. H., Marone, D., Petruzzino, G., Corneti, S., Ormanbekova, D., … Bassi, F. M. (2020). The global durum wheat panel (GDP): An international platform to identify and exchange beneficial alleles. *Frontiers in Plant Science, 11*, Article 569905. https://doi.org/10.3389/fpls.2020.569905

Miazzi, M. M., Babay, E., De Vita, P., Montemurro, C., Chaabane, R., Taranto, F., & Mangini, G. (2022). Comparative genetic analysis of durum wheat landraces and cultivars widespread in Tunisia. *Frontiers in Plant Science, 13*, Article 939609. https://doi.org/10.3389/fpls.2022.939609

Morcia, C., Terzi, V., Ghizzoni, R., Vaiuso, C., Delogu, C., Andreani, L., Venturini, A., Carnevali, P., Pompa, P. P., & Tumino, G. (2021). Digital PCR for genotype quantification: A case study in a pasta production chain. *Biology, 10*(5), 419. https://doi.org/10.3390/biology10050419

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

Palermo, N., Buffagni, V., Vurro, F., Impollonia, G., Pignone, D., Janni, M., Nguyen, H. T., Dembech, E., & Marmiroli, N. (2024). From landraces to haplotypes, exploiting a genomic and phenomic approach to identify heat tolerant genotypes within durum wheat landraces. *Environmental and Experimental Botany, 228*, Article 105986. https://doi.org/10.1016/j.envexpbot.2024.105986

Palombieri, S., Bonarrigo, M., Potestio, S., Sestili, F., Messina, B., Russo, G., Miceli, C., Frangipane, B., Genduso, M., Delogu, C., Andreani, L., & Masci, S. (2024). Characterization among and within Sicilian tetraploid wheat landraces by gluten protein analysis for traceability purposes. https://doi.org/10.3390/plants13050741.

Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics, 35*(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pasqualone, A., Alba, V., Mangini, G., Blanco, A., & Montemurro, C. (2010). Durum wheat cultivar traceability in PDO altamura bread by analysis of DNA microsatellites. *European Food Research and Technology, 230*(5), 723–729. https://doi.org/10.1007/s00217-009-1210-1

Perry, D. J. (2004). Identification of Canadian durum wheat varieties using a single PCR. *Theoretical and Applied Genetics, 109*(1), 55–61. https://doi.org/10.1007/s00122-004-1597-9

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Ramirez-Gonzalez, R. H., Uauy, C., & Caccamo, M. (2015). PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics, 31*(12), 2038–2039. https://doi.org/10.1093/bioinformatics/btv069

Rimbert, H., Darrier, B., Navarro, J., Kitt, J., Choulet, F., Leveugle, M., Duarte, J., Rivière, N., Eversole, K., on behalf of The International Wheat Genome Sequencing Consortium, Le Gouis, J., on behalf The BreedWheat Consortium, Davassi, A., Balfourier, F., Le Paslier, M.-C., Berard, A., Brunel, D., Feuillet, C., Poncet, C., … Paux, E. (2018). High throughput SNP discovery and genotyping in hexaploid wheat. *PLoS One, 13*(1), Article e0186329. https://doi.org/10.1371/journal.pone.0186329

Roncallo, P. F., Larsen, A. O., Achilli, A. L., Pierre, C. S., Gallo, C. A., Dreisigacker, S., & Echenique, V. (2021). Linkage disequilibrium patterns, population structure and diversity analysis in a worldwide durum wheat collection including Argentinian genotypes. *BMC Genomics, 22*(1), 233. https://doi.org/10.1186/s12864-021-07519-z

Ruisi, P., Frangipane, B., Amato, G., Frenda, A. S., Plaia, A., Giambalvo, D., & Saia, S. (2015). Nitrogen uptake and nitrogen fertilizer recovery in old and modern wheat genotypes grown in the presence or absence of interspecific competition. *Frontiers in Plant Science, 6*. https://doi.org/10.3389/fpls.2015.00185

Sales, E., García-Romeral, J., & Domingo, C. (2023). Bioinformatics approach for developing a minimum set of SNP markers for identification of temperate japonica rice varieties cultivated in Spain. *PLoS One, 18*(6), Article e0286839. https://doi.org/10.1371/journal.pone.0286839

Scandurra, A., Corinzia, S. A., Caruso, P., Cosentino, S. L., & Testa, G. (2024). Productivity of wheat landraces in rainfed and irrigated conditions under conventional and organic input in a semiarid mediterranean environment. *Agronomy, 14*(10), 2338. https://doi.org/10.3390/agronomy14102338

Scavo, A., Pandino, G., Restuccia, A., Caruso, P., Lombardo, S., & Mauromicale, G. (2022). Allelopathy in durum wheat landraces as affected by genotype and plant part. *Plants, 11*(8), 1021. https://doi.org/10.3390/plants11081021

Sertse, D., You, F. M., Klymiuk, V., Haile, J. K., N'Diaye, A., Pozniak, C. J., Cloutier, S., & Kagale, S. (2023). Historical selection, adaptation signatures, and ambiguity of introgressions in wheat. *International Journal of Molecular Sciences, 24*(9), 8390. https://doi.org/10.3390/ijms24098390

Shewry, P. R., Faulks, A. J., Pratt, H. M., & Miflin, B. J. (1978). The varietal identification of single seeds of wheat by sodium dodecyl-sulphate polyacrylamide gel electrophoresis of gliadin. *Journal of the Science of Food and Agriculture, 29*(10), 847–849. https://doi.org/10.1002/jsfa.2740291004

Tang, D., Chen, M., Huang, X., Zhang, G., Zeng, L., Zhang, G., Wu, S., & Wang, Y. (2023). SRplot: A free online platform for data visualization and graphing. *PLoS One, 18*(11), Article e0294236. https://doi.org/10.1371/journal.pone.0294236

Tang, W., Lin, J., Wang, Y., An, H., Chen, H., Pan, G., Zhang, S., Guo, B., Yu, K., Li, H., Fang, X., & Zhang, Y. (2022). Selection and validation of 48 KASP markers for variety identification and breeding guidance in conventional and hybrid rice (*Oryza sativa* L.). *Rice, 15*(1), 48. https://doi.org/10.1186/s12284-022-00594-0

Taranto, F., D'Agostino, N., Rodriguez, M., Pavan, S., Minervini, A. P., Pecchioni, N., Papa, R., & De Vita, P. (2020). Whole genome scan reveals molecular signatures of divergence and selection related to important traits in durum wheat germplasm. *Frontiers in Genetics, 11*. https://doi.org/10.3389/fgene.2020.00217

Taranto, F., Di Serio, E., Miazzi, M. M., Pavan, S., Saia, S., De Vita, P., & D'Agostino, N. (2022). 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, 12*(6), 1326. https://doi.org/10.3390/agronomy12061326

Thioulouse, J., Dray, S., Dufour, A.-B., Siberchicot, A., Jombart, T., & Pavoine, S. (2018). *Multivariate analysis of ecological data with ade4*. Springer. https://doi.org/10.1007/978-1-4939-8850-1

Varia, F., Macaluso, D., Vaccaro, A., Caruso, P., & Guccione, G. D. (2021). The adoption of landraces of durum wheat in Sicilian organic cereal farming analysed using a system dynamics approach. *Agronomy, 11*(2), 319. https://doi.org/10.3390/agronomy11020319

Velimirović, A., Jovović, Z., Perović, D., Lehnert, H., Mikić, S., Mandić, D., Pržulj, N., Mangini, G., & Finetti-Sialer, M. M. (2023). SNP diversity and genetic structure of "Rogosija", an old Western balkan durum wheat collection. *Plants, 12*(5), 1157. https://doi.org/10.3390/plants12051157

Visioli, G., Giannelli, G., Agrimonti, C., Spina, A., & Pasini, G. (2021). Traceability of Sicilian durum wheat landraces and historical varieties by high molecular weight glutenins footprint. *Agronomy, 11*(1). https://doi.org/10.3390/agronomy11010143

Wang, J. (2024). PopCluster: A population genetics model-based toolset for simulating, inferring and visualising individual admixture and population structure. *Molecular Ecology Resources*. , Article e14058. https://doi.org/10.1111/1755-0998.14058

Wang, P., Tian, T., Ma, J., Liu, Y., Zhang, P., Chen, T., Shahinnia, F., & Yang, D. (2022). Genome-wide association study of kernel traits using a 35K SNP array in bread wheat (Triticum aestivum L.). *Frontiers in Plant Science, 13*. https://doi.org/10.3389/fpls.2022.905660

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., Maccaferri, M., Salvi, S., Milner, S. G., Cattivelli, L., Mastrangelo, A. M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., Lillemo, M., Mather, D., Appels, R., … Akhunov, E. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal, 12*(6), 787–796. https://doi.org/10.1111/pbi.12183

Wei, T., & Simko, V. (2024). R package "corrplot": Visualization of a correlation matrix. https://github.com/taiyun/corrplot.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. https://doi.org/10.1007/978-0-387-98141-3

Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, P. A., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., Pirani, A., Webster, T., Brew, F., Bloor, C., King, J., West, C., Griffiths, S., King, I., Bentley, A. R., & Edwards, K. J. (2016). High-density SNP genotyping array for

hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnology Journal, 14*(5), 1195–1206. https://doi.org/10.1111/pbi.12485

Winfield, M., Burridge, A., Ordidge, M., Harper, H., Wilkinson, P., Thorogood, D., Copas, L., Edwards, K., & Barker, G. (2020). Development of a minimal KASP marker panel for distinguishing genotypes in apple collections. *PLoS One, 15*(11), Article e0242940. https://doi.org/10.1371/journal.pone.0242940

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). Ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution, 8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628