



OPEN ACCESS

EDITED BY

Ahmed Sallam,
Leibniz Institute of Plant Genetics and Crop
Plant Research (IPK), Germany

REVIEWED BY

Shamseldeen Shehabeldin Eltaher,
University of Sadat City, Egypt
Clizia Villano,
University of Naples Federico II Italy

*CORRESPONDENCE

Hilal Betül Kaya
✉ hilalbetul.kaya@cbu.edu.tr

RECEIVED 12 December 2022

ACCEPTED 06 April 2023

PUBLISHED 10 May 2023

CITATION

Kaya HB, Dilli Y, Oncu-Oner T and Ünal A
(2023) Exploring genetic diversity and
population structure of a large grapevine
(*Vitis vinifera* L.) germplasm collection
in Türkiye.
Front. Plant Sci. 14:1121811.
doi: 10.3389/fpls.2023.1121811

COPYRIGHT

© 2023 Kaya, Dilli, Oncu-Oner and Ünal.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Exploring genetic diversity and population structure of a large grapevine (*Vitis vinifera* L.) germplasm collection in Türkiye

Hilal Betül Kaya^{1*}, Yıldız Dilli², Tulay Oncu-Oner¹
and Akay Ünal²

¹Department of Bioengineering, Manisa Celal Bayar University, Manisa, Türkiye, ²Republic of Türkiye Ministry of Agriculture and Forestry, Viticulture Research Institute, Manisa, Türkiye

Grapevine (*Vitis Vinifera* L.) has been one of the significant perennial crops in widespread temperate climate regions since its domestication around 6000 years ago. Grapevine and its products, particularly wine, table grapes, and raisins, have significant economic importance not only in grapevine-growing countries but also worldwide. Grapevine cultivation in Türkiye dates back to ancient times, and Anatolia is considered one of the main grapevine migration routes around the Mediterranean basin. Turkish germplasm collection, conserved at the Turkish Viticulture Research Institutes, includes cultivars and wild relatives mainly collected in Türkiye, breeding lines, rootstock varieties, and mutants, but also cultivars of international origin. Genotyping with high-throughput markers enables the investigation of genetic diversity, population structure, and linkage disequilibrium, which are crucial for applying genomic-assisted breeding. Here, we present the results of a high-throughput genotyping-by-sequencing (GBS) study of 341 genotypes from grapevine germplasm collection at Manisa Viticulture Research Institute. A total of 272,962 high-quality single nucleotide polymorphisms (SNP) markers on the nineteen chromosomes were identified using genotyping-by-sequencing (GBS) technology. The high-density coverage of SNPs resulted in an average of 14,366 markers per chromosome, an average polymorphism information content (PIC) value of 0.23 and an expected heterozygosity (He) value of 0.28 indicating the genetic diversity within 341 genotypes. LD decayed very fast when r^2 was between 0.45 and 0.2 and became flat when r^2 was 0.05. The average LD decay for the entire genome was 30 kb when $r^2 = 0.2$. The PCA and structure analysis did not distinguish the grapevine genotypes based on different origins, highlighting the occurrence of gene flow and a high amount of admixture. Analysis of molecular variance (AMOVA) results indicated a high level of genetic differentiation within populations, while variation among populations was extremely low. This study provides comprehensive information on the genetic diversity and population structure of Turkish grapevine genotypes.

KEYWORDS

Vitis Vinifera L., genetic characterization, linkage disequilibrium, divergent loci, genotyping by sequencing

Introduction

Grapevine (*Vitis vinifera* L.) is one of the most important fruit crops worldwide (Jaillon et al., 2007). Besides its broad uses in wine industries, this fruit crop has been used to produce raisins, fresh fruits, juices, and leaves. In addition, grapes are an important source of secondary metabolites used in the cosmetic, food, and pharmaceutical industries (Ali et al., 2010). The beneficial effects of these secondary metabolites on several diseases have been reported in some studies (Folts, 2002; Tyagi et al., 2003; Tenkumo et al., 2020). In 2020, the total area of harvested grapes worldwide was 7 million hectares, with 78 million tons of grapes (FAOSTAT, 2018). China is the largest producer of grapes in the world, followed by Italy, the USA, Spain, and France. Türkiye is a critical producer in the grape market, ranked 6th in the world in 2020 according to total grape production.

The archaeological record suggests that grapevine domestication began in southeastern Anatolia (the Asian part of Türkiye) or the region known as Transcaucasian about 6000 to 7000 years ago before grapevine cultivars spread throughout neighboring regions of Europe and Northern Africa (Zohary et al., 2012; Mc Govern, 2003). During the spreading period *via* different routes, high genetic diversity among grapevine cultivars is a result of sexual reproduction (Vouillamoz and Grando, 2006), (Myles et al., 2011), and spontaneous mutations (Margaryan et al., 2021). According to *Vitis* international variety catalog (Maul and Töpfer, 2015), there are more than 12,000 documented grapevine cultivars. Still, only a small number of cultivars have economic importance worldwide (Bouby et al., 2013), and growers have been mainly focused on them. Although the rapid spread of grapevine disease phyloxera has caused a significant decrease in grapevine genetic diversity, the elimination of old and local cultivars from vineyards due to their low-quality characteristics over time has also led to a gradual reduction in genetic resources (Tello et al., 2019a).

Grapevine germplasm collections, including cultivars, wild relatives, breeding lines, rootstock varieties, and mutants, have been constructed for characterization, conservation, use, and development of grapevine genetic resources in the leading grapevine-growing countries such as Spain, France, Germany, Italy, and the USA (Martín et al., 2006; Lacombe et al., 2007; Cipriani et al., 2010; Maul et al., 2012; Žulj Mihaljević et al., 2020). In addition, these germplasm collections have been prevalently used in molecular studies, including identification (Laucou et al., 2011; Laucou et al., 2018), molecular characterization (Lopes et al., 1999; Nicolas et al., 2016), and mapping studies (Myles et al., 2011; Laucou et al., 2018; Guo et al., 2019). In Türkiye, there are two Viticulture Research Institutes in Tekirdağ and Manisa, Tekirdağ Viticulture Research Institute hold approximately 1200 genotypes of cultivars and some important breeding lines with wild relatives (Söylemezoğlu et al., 2016). With the efforts of protection of Turkish Germplasm collection, Manisa Viticulture Research Institutes was established in 1930 and started to conduct surveys to preserve Turkish grapes in the Aegean region. In the following years, some other important genotypes, including national and international cultivars, were transferred from Ege University and Aegean Agricultural Research

Institute, one of the government research institutes of TAGEM (General Directorate of Agricultural Research and Policies).

Türkiye is one of the most favorable locations for viticulture in terms of ecological conditions, and grape production ranks first in fruit production, making a significant contribution to Türkiye's economy. However, grapevine cultivation is mainly concentrated in the western part of the country, especially in the Aegean region. Mediterranean and Marmara regions rank second and third, respectively, based on both areas of vineyards and grape production. Although vineyards are spread throughout the country, the other regions cover a relatively small area for grapevine production. In Türkiye, Manisa is the leading grape producer city, with 809 thousand hectares of vineyard areas, accounting for approximately 38% of the total output in the country (TUİK, 2019). In the exploration survey of grapevine genetic resources of Türkiye in 2021, 1439 different genotypes were protected in the Turkish National Grapevine Genetic Reserve in Tekirdağ Viticulture Research Institute.

Characterizing the genetic diversity and population structure of grapevine germplasms is crucial to provide valuable genetic resources for grapevine genetic improvement and to protect and to develop this vital crop (Emanuelli et al., 2013; De Oliveira et al., 2020). Unfortunately, mislabeling of genotypes as homonyms or synonyms is widespread among grapevine genotypes, leading to problems in classification. For a sustainable viticulture industry, extensive genetic diversity should be identified and maintained to be able to develop new grape cultivars with desired traits through marker-assisted breeding (Di Gaspero and Cattonaro, 2010; Myles et al., 2011).

Ampelographic identification, including the morphological characteristics of leaves, shoot tips, branches, fruit bunches, and berries, has been traditionally used to characterize grapevine cultivars (Topfer et al., 2009). However, the molecular characterization of plant germplasm is the preferred approach to identify genetic diversity and variation within the population (Gago et al., 2022). Information regarding genetic diversity, population structure, and gene flow are essential for accelerating the development of efficient breeding strategies, and DNA markers are powerful tools for identifying and characterizing diverse genotypes (Myles et al., 2011; De Oliveira et al., 2020). In addition, DNA markers provide useful information in theoretical and applied research fields for grapevine breeding, such as cultivar identification, determination of genetic diversity, paternity analyses, characterization of large grapevine germplasms, linkage map construction, and detection of marker-trait associations *via* QTL (Quantitative Trait Locus) and association mapping (Tomić et al., 2013). Although various types of molecular markers, including RFLP (Restriction fragment length polymorphism), RAPD (Random Amplified Polymorphic DNA), SSR (Simple Sequence Repeats), and SNP (Single Nucleotide Polymorphism), have been utilized in grapevine (Lodhi et al., 1995; Dalbó et al., 2000; Adam-Blondon et al., 2004), SSRs have been extensively used because of their high polymorphism, codominant nature, and reproducibility (Riaz et al., 2004; This et al., 2004). However, using SSR as a genotyping platform for large-scale germplasm screening has some

drawbacks, including being time-consuming, labor-intensive, and low-throughput (Deschamps et al., 2012).

Next-generation sequencing (NGS) has enabled the discovery of thousands of markers in large and diverse germplasm collections (Cossa et al., 2013). NGS-based genome-wide SNP (single nucleotide polymorphism) markers have emerged as a powerful molecular tool for germplasm characterization and population structure studies. Genotyping by sequencing (GBS) is a simple and relatively inexpensive technique that reduces genome complexity using restriction enzymes for NGS-based SNP identification. With this advantage, GBS has been implemented in various woody perennial species such as olive (D'Agostino et al., 2018; Kaya et al., 2019), peach (Bielenberg et al., 2015), apple (Gardner et al., 2014), sweet cheery (Guajardo et al., 2015), almond (Goonetilleke et al., 2018), and oil palm (Teh et al., 2016), and found to be effective for high-throughput genotyping. The first report on the use of GBS in grapevine was to discover SNPs in an F1 segregation population to analyze the inheritance of powdery mildew resistance (Barba et al., 2014). However, the same group reports high rates of missingness and heterozygote under-calling. Because of the highly heterozygous and diverse grapevine genome, they developed a modular approach to overcome these problems. Recently, GBS was applied for clarification of evolutionary relationships among North American *Vitis* species (Klein et al., 2018), characterization of some accessions in the USDA (United States Department of Agriculture) *Vitis* germplasm collections (Klein et al., 2018), and identification of marker-trait associations in grapevine by GWAS (Genome-wide association studies) (Guo et al., 2019; Flutre et al., 2020) and QTL mapping (Barba et al., 2014) which demonstrated the suitability of GBS for the high-throughput genotyping in grapevine. RAD (Restriction-site Associated DNA)-sequencing, a similar approach to GBS, has also been implemented to characterize the relatedness between wild and cultivated grapevine in a germplasm collection (Marrano et al., 2017). The present study is the first report on discovering GBS-generated SNP markers in a diverse collection of Turkish grapevine genotypes. In this study, we aimed (1) to characterize 341 grapevine genotypes originating from different regions of Turkey, mostly Aegean Region genotypes, and some other countries, (2) to detect synonymies, homonymies, and misnaming and, (3) to assess the level of genetic diversity. The results not only indicate the nature and extent of the genetic diversity in the grapevine genotypes but also estimate population structure and characterize the LD pattern.

Materials and methods

Plant materials

A total of 341 genotypes from grapevine germplasm collection in Manisa Viticulture Research Institute (38°N, 27°E) were genetically characterized in this study. This grapevine sample set includes genotypes from Aegean Region, Marmara Region, Mediterranean Region, Central Anatolia, Eastern Anatolia Region, Black Sea Region, Southeastern Anatolia Region as well as 80 genotypes from unknown locations in Türkiye and nine

genotypes from France, nine genotypes from Italy, one genotype from Spain, one genotype from Greece, one genotype from Uzbekistan, three genotypes from the USA and two unknown genotypes. A list of all samples and their specific codes is available in Supplementary Table S1.

DNA extraction, library preparation, and sequencing

DNA was extracted from young leaves using the CTAB (Cetyltrimethylammonium bromide) protocol (Lodhi et al., 1994). The quality and quantity of genomic DNA were evaluated using the Qubit™ Fluorometer 3.0 (Invitrogen). DNA concentrations were adjusted to 20 ng/ul and used for GBS library preparation. GBS libraries were prepared by digestion of DNA with the *ApeKI* restriction enzyme in 96-plex, where each plate included a single random blank well. PCR amplification was performed to generate the GBS libraries, and DNA was sequenced on a Genome Analyzer II device in a single flowcell channel (Illumina Inc., USA). GBS was carried out at the University of Wisconsin-Madison Biotechnology Center as previously described (Elshire et al., 2011).

SNP calling, filtering, and imputation

Raw reads were trimmed to remove any sequencing adapters and low-quality bases by trimming software skewer (Jiang et al., 2014). Trimmed reads were aligned to the 12x.0 version of the PN40024 reference genome (Jaillon et al., 2007; Adam-Blondon et al., 2011) using Bowtie 2 (Langmead and Salzberg, 2012), and TASSEL reference-based GBS analysis pipeline was used for SNP calling (Glaubitz et al., 2014). SNPs with minor allele frequency (MAF) > 0.05 and missing data > 0.80 at the markers and genotypes level were filtered, and monomorphic SNPs were removed. Imputation was carried out using Beagle 4.1 (Browning and Browning, 2016) with a probability > 0.80. The minor allele frequency (MAF) was calculated using TASSEL (v5.2.64) (Bradbury et al., 2007). Finally, the polymorphism information content (PIC) values were calculated using SNP data using the following equation (Botstein et al., 1980). P_i and P_j show the population frequency of the i th and j th allele.

$$PIC = 1 - \sum_{i=1}^n P_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2P_i^2 P_j^2$$

Observed (H_o) and expected (H_e) heterozygosity were calculated in the “snpReady” package (Granato et al., 2018) in R. Distributions of SNPs within 1Mb window size after filtering were visualized using the R package “CMplot” (Yin et al., 2021).

Analysis of the genetic diversity

TASSEL (v5.2.64) (Bradbury et al., 2007) was used to calculate the kinship (centered IBS) matrix and genetic distance matrix

between all 341 grapevine genotypes. Kinship heatmap and histogram were visualized using the heatmap.2 function in R package “gplots” (Warnes et al., 2009). The phylogenetic analysis was performed using the neighbor-joining method implemented in TASSEL (v5.2.64) (Bradbury et al., 2007). The phylogenetic tree was saved in Newick format in TASSEL (v5.2.64) (Bradbury et al., 2007) and visualized using the iTOL v4.3.3 online tool (Letunic and Bork, 2019).

Analysis of population structure

A principal component analysis (PCA) was conducted using the “prcomp” function in R (Team, R.C., 2016). The PCA plots (PC1 vs. PC2, PC2 vs. PC3, PC1 vs. PC3) were constructed using the “ggplot2” package (Wickham, 2011) in R.

Population structure was investigated to infer the most likely number of ancestral populations using ADMIXTURE software version 1.3.0 (Alexander et al., 2009). Using default settings, analysis was carried out with K (number of populations) ranging from 1 to 14. The optimal number of K values was determined based on the lowest cross-validation (CV) error program reported. The admixture proportions of each genotype were visualized using the “ggplot2” package (Wickham, 2011) in R, and genotypes were assigned to a specific cluster when the estimated membership coefficient was above 0.7. Pairwise F_{st} (fixation index) values were calculated using ADMIXTURE software (Alexander et al., 2009). Pairwise F_{st} values for each SNP between the populations identified by ADMIXTURE software were also estimated using the Weir and Cockerham algorithm (Weir and Cockerham, 1984) in VCFtools 0.1.16 (Danecek et al., 2011) to identify the divergent loci. Distribution of the pairwise F_{st} values of each SNP across the 19 grapevine chromosomes was displayed with “ggplot2” package (Wickham, 2011) in R. SNP loci with $F_{st} \geq 0.5$ were identified as divergent, and genes associated with divergent SNPs were found based on the PN40024 reference genome using the Jbrowse feature of Phytozome v.13 (<http://phytozome.jgi.doe.gov/pz/portal.html>).

Analysis of molecular variance (AMOVA) was performed to investigate the genetic differentiation within and among populations based on the defined populations by STRUCTURE and geographic origin of the genotypes using the “poppr” package in R (Kamvar et al., 2014).

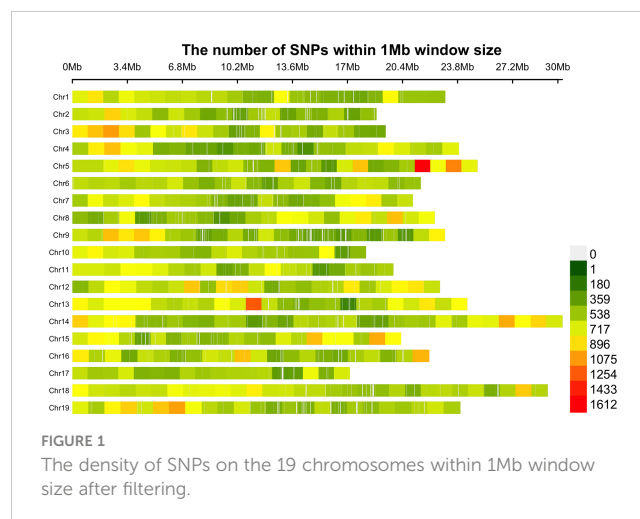
Linkage disequilibrium analysis

The level of LD was evaluated based on the squared allele frequency (r^2) between SNPs on each chromosome for the entire germplasm collection using TASSEL (Bradbury et al., 2007). The SNPs were considered to be in significant LD when $P < 0.01$. The trend in LD decay was investigated by plotting pairwise LD values (r^2) within each 300-kb bin against the distance between SNPs for each chromosome and the whole genome in R (Team, R.C., 2016). It was fitted by nonlinear regression using the expectation of r^2 between adjacent sites proposed by Hill and Weir (Hill and Weir, 1988).

Results

SNP genotyping and distribution of SNPs in the grapevine genome

GBS sequencing of 341 grapevine genotypes generated >1 billion reads, with an average of 3,056,208 reads per sample. A set of 470,672 unfiltered SNPs was obtained from these data and 36,464 SNPs were mapped to the scaffolds not yet assigned to specific chromosomes. After filtering for minor allele frequency threshold of 0.05, >20% missing, and monomorphic SNPs, 272,962 genome-wide SNP markers were obtained and used for downstream analysis. The number of SNPs within the 1Mb window size after filtering for each chromosome is shown in Figure 1. The high-density coverage of SNPs resulted in an average of 14,366 markers per chromosome, ranging from 10,063 on chromosome 17 to 19,609 on chromosome 18. All the SNPs were mapped onto the 19 chromosomes covering a total of 425.8 Mb of the grapevine genome. As shown in Table S2, the chromosome size covered by SNPs ranged from 17.1 Mb (chromosome 17) to 30.3 Mb (chromosome 14). Individually, the average number of SNPs per Mb varied from 570.6 on chromosome 2 to 720.6 on chromosome 5. The average minor allele frequency was 0.196 (Table S2), and 85.8% of all the SNPs had $MAF > 5\%$ (Figure 2). The polymorphic information content (PIC) values ranged from 0.0087 to 0.38, with an average of 0.23 (Figure 2). PIC values across 19 chromosomes were very similar, with a mean of 0.23 (Supplementary Table S2). The observed heterozygosity (H_o) values varied between 0.00 to 0.99, with an average of 0.23. Overall mean expected heterozygosity was higher than the observed heterozygosity. The expected heterozygosity (H_e) values ranged from 0.08 to 0.50, with an overall mean of 0.28 (Figure 2). All the raw sequencing reads for all genotypes have been submitted to the NCBI Sequence Read Archive and deposited under the “BioProject ID”: PRJNA742054.



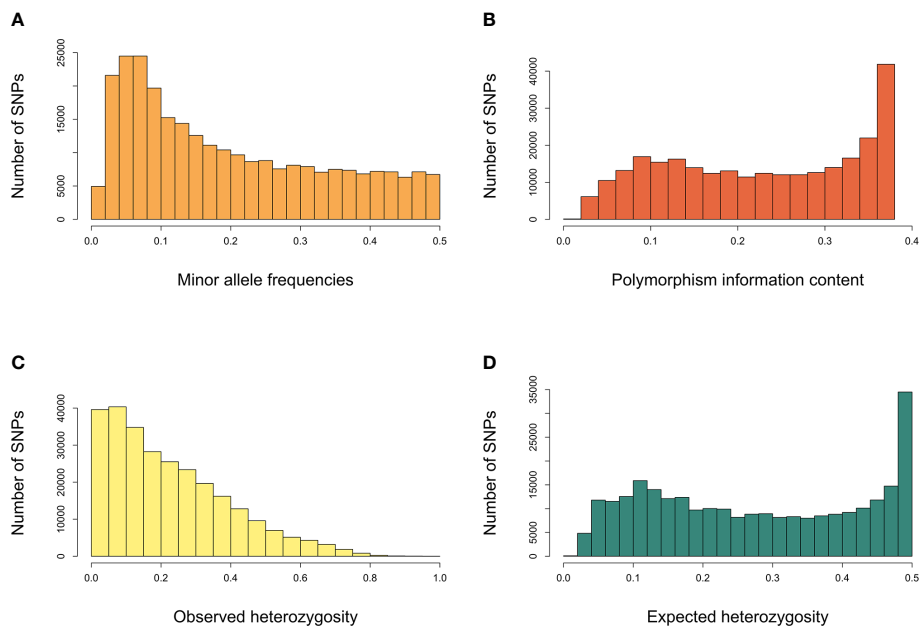


FIGURE 2

Characteristic statistics of SNPs. (A) minor allele frequencies, (B) polymorphism information content, (C) observed heterozygosity (H_o), (D) expected heterozygosity (H_e).

Genetic diversity

A total of 272,962 SNPs on the nineteen chromosomes were used to evaluate genetic diversity in 341 genotypes. Different complementary approaches were used to estimate the genetic diversity of grapevine germplasm collection. We calculated marker-based kinship coefficients between pairs of 341 genotypes which ranged from 0.00 to 1.78 with a mean value of 0.02 (Supplementary Table S3). Pairwise relative kinship values of 0 accounted for 67% of all kinship coefficients. Figure 3A shows the distribution of the relative kinship values for all marker pairs.

The genetic diversity among grapevine genotypes was also shown by the IBS-based genetic distance matrix. The distribution of genetic distances is shown in Figure 3B. The genetic distance of pairwise comparisons of the grapevine genotypes varied from 0.138 to 0.325 with an average dissimilarity of 0.277 (Supplementary

Table S4). Most of the genetic distances were between 0.25 and 0.30. The smallest genetic distance (0.138) was observed between “İnek Memesi” (c-4-3) and “Çavuş Malaga” (a-4-5). The maximum genetic distance (0.324) was found between “Kırlı kadın parmağı” (a-9-3) and “Merlot” (m). For the three tested reference genotypes, the mean genetic distance was 0.301, 0.279, and 0.271 for “Merlot” (m), “Cabernet Sauvignon” (cs), and “Cardinal” (c), respectively.

A neighbor-joining based cluster analysis was implemented to gain further insight into the genetic diversity among grapevine genotypes. The 341 grapevine genotypes were classified into seven major clusters, each containing genotypes from different regions, indicating a high amount of genetic admixture between genotypes (Figure 4).

The most significant number of genotypes convened in cluster 1, containing 83 genotypes with an average genetic distance of 0.268. Cluster 1 was divided into two sub-clusters, A and B. Sub-

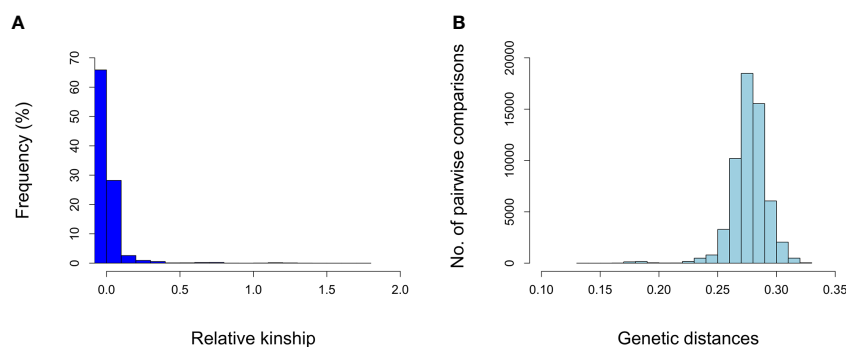
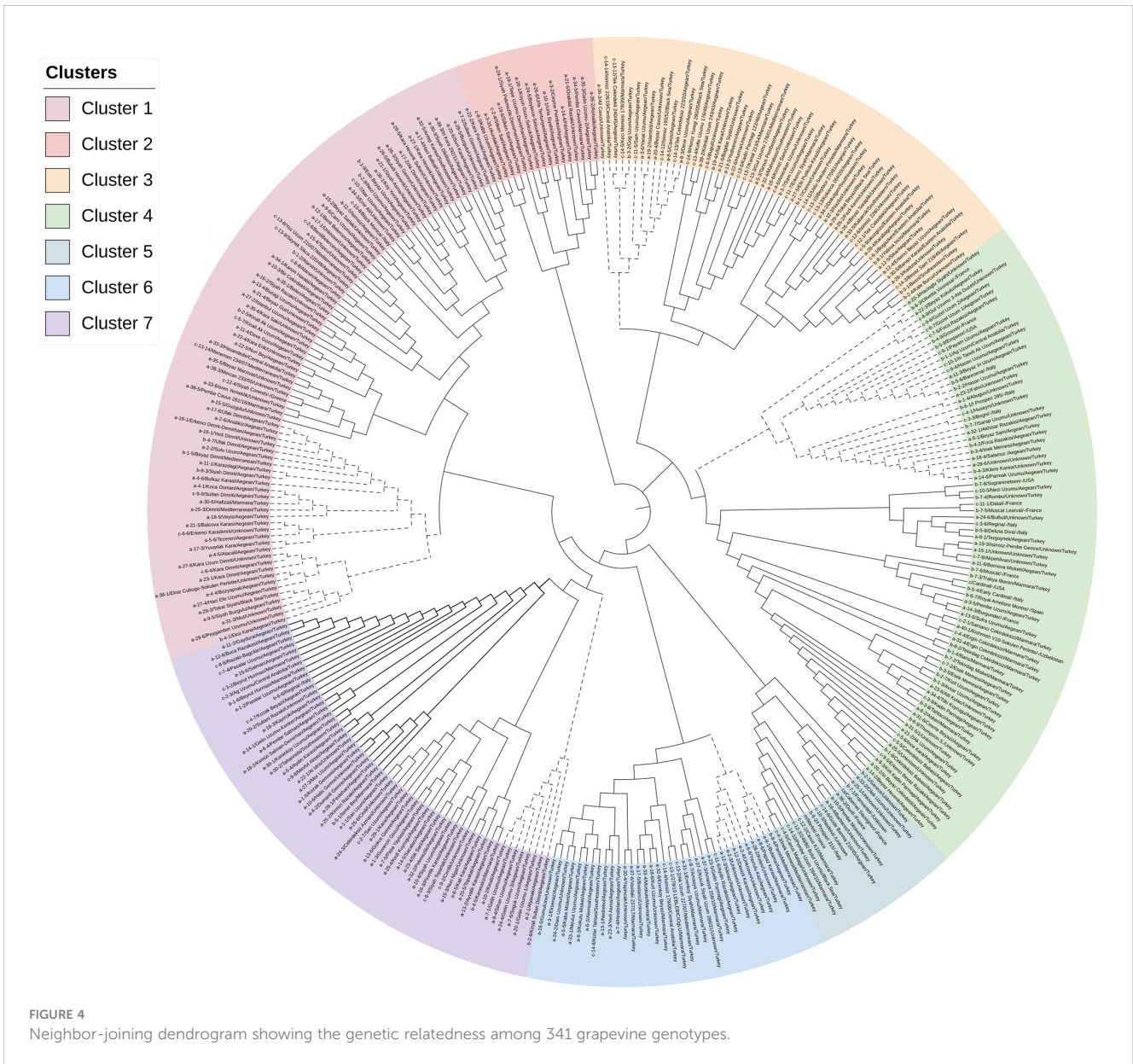


FIGURE 3

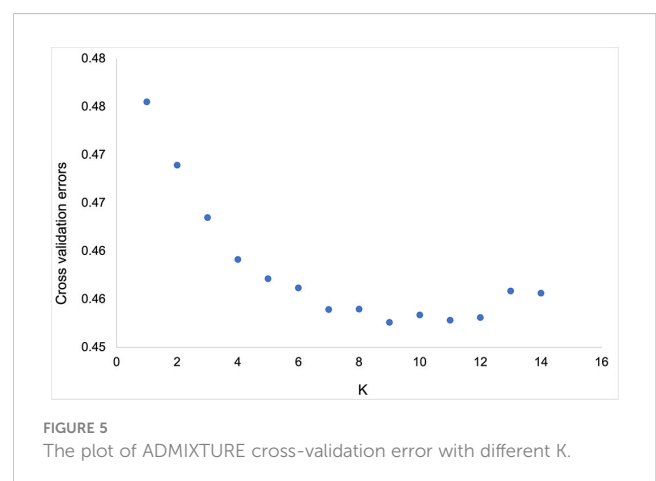
Histograms showing the distribution of pairwise relative kinship values (A) and distribution of pairwise genetic distances (B).

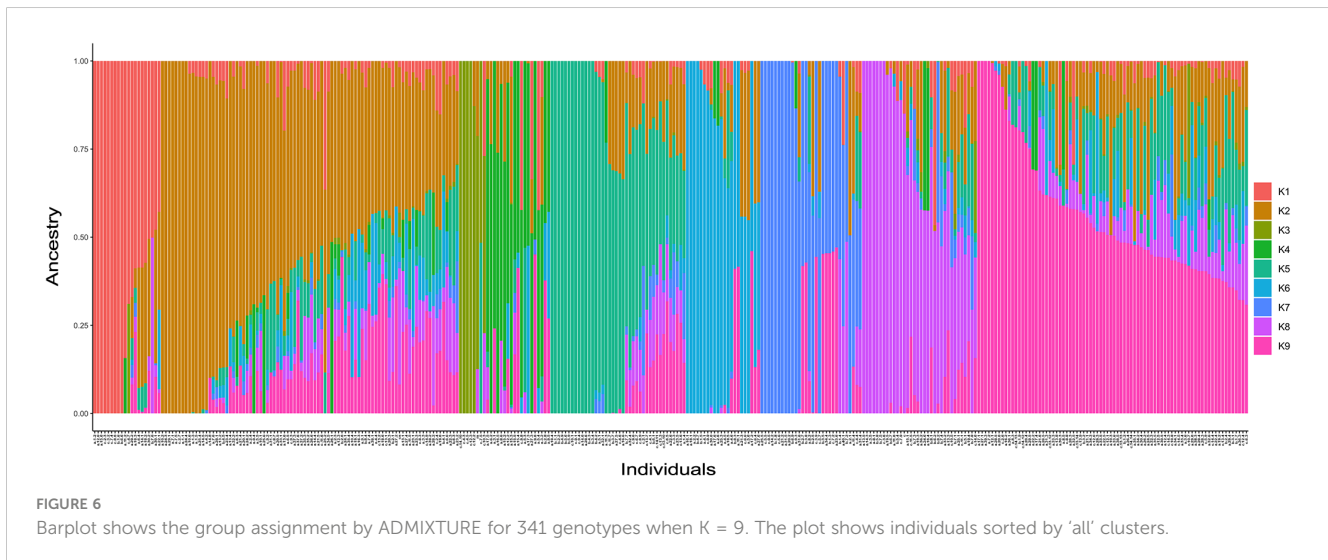


cluster 1A included 27 genotypes from the Aegean Region, two genotypes from the Mediterranean Region, one each from the Black Sea Region, Southeastern Region, Central Anatolia Region, Italy, Greece, and 24 undefined genotypes. Sub-cluster 1B contained 31 genotypes: twenty-one from the Aegean region, two from the Mediterranean Region, one from the Marmara and the Black Sea Region, and 18 from the unknown location.

Cluster 2 (mean genetic distance=0.261) consisted of 16 genotypes, 14 from the Aegean Region, one from the Marmara region, and one from an unknown location.

Cluster 3 (mean genetic distance=0.262) contained 54 genotypes subdivided into two sub-clusters (3A and 3B). Sub-cluster 3A (n=13) had eight genotypes from the Aegean Region and one each from the Marmara Region, Central Anatolia, and the Black Sea Regions. It also contained one genotype from an unknown location. Sub-cluster 3B included 41 genotypes:





fourteen from the Aegean region, four from each of the Marmara and the Southeastern Regions, three from the Eastern Anatolia Region, two from the Black Sea Region, and one from the Mediterranean Region. There were also thirteen genotypes from unknown locations in this sub-cluster.

Cluster 4 (average genetic distance=0.266) is subdivided into two sub-clusters: sub-cluster 4A and sub-cluster 4B, which consisted of 30 and 49 genotypes, respectively. Sub-cluster 4A included 30 genotypes comprising 15 genotypes from the Aegean Region, three from Italy, two from France, and one from the Central Anatolia Region and the USA. It included eight genotypes from an unknown location. Sub-cluster 4B included 49 genotypes that contained “Cardinal” (genotype c), one of the three reference grapevine genotypes. Among 49 genotypes, twenty-one were from the Aegean Region, eight from the Marmara Region, four from France, three from Italy, two from the USA, and one from Uzbekistan and Spain. Nine unknown genotypes did not have origin information.

Cluster 5 included 17 genotypes composed of 6 from an unknown location in Türkiye, two from unknown countries, three from France, two from Marmara, and one each genotype from the Aegean Region, Mediterranean Region, Black Sea Region, and Italy.

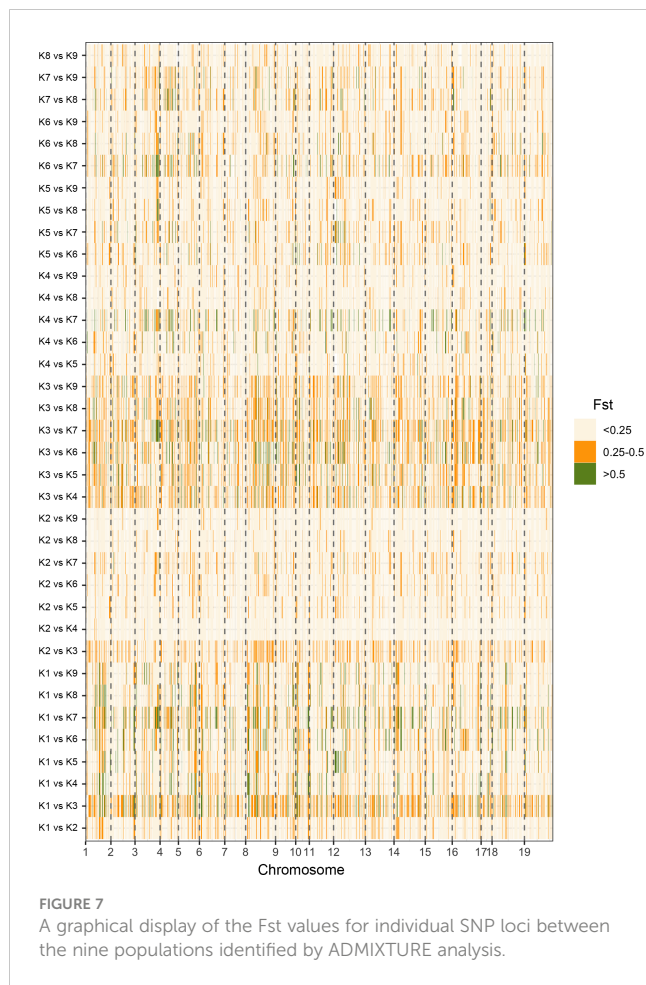
These genotypes were more genetically distinct than other clusters, with an average genetic distance of 0.273. “İnek Memesi” (genotype c-4-3) and “Çavuş Malaga” (genotype a-4-5) had the lowest genetic distance (0.138) values and were placed in this cluster. In addition, “Merlot” (m) and “Cabernet Sauvignon” (cs) reference grapevine genotypes were placed in cluster 5.

Cluster 6 consisted of 35 genotypes from 5 different geographic regions of Türkiye, subdivided into two sub-clusters (6A and 6B). While sub-cluster 6A included nine genotypes from the Aegean, Marmara, and Mediterranean Regions, sub-cluster 6B consisted of 26 genotypes, including ten genotypes from the Aegean Region, seven from the Marmara, one each from the Mediterranean, Central Anatolia and Southeastern regions. Among 35 genotypes in this cluster, there were three and six from unknown locations.

There were 57 genotypes in Cluster 7 with an average genetic distance of 0.262, separated into three sub-clusters. Sub-cluster 7A included eight genotypes which were all from the Aegean Region. Sub-cluster 7B contained 21 genotypes, of which 15 were from the Aegean Region, one from the Marmara, and five from an unknown location. Sub-cluster 7C contained 28 genotypes and most of them were from the Aegean Region (n=19) and 3 genotypes from the

TABLE 1 Fst values between estimated populations.

Cluster	K1	K2	K3	K4	K5	K6	K7	K8	K9
K1	0								
K2	0.122	0							
K3	0.255	0.195	0						
K4	0.173	0.103	0.24	0					
K5	0.148	0.091	0.212	0.136	0				
K6	0.171	0.11	0.236	0.145	0.127	0			
K7	0.193	0.135	0.251	0.181	0.135	0.17	0		
K8	0.144	0.079	0.205	0.13	0.097	0.125	0.145	0	
K9	0.137	0.077	0.202	0.115	0.088	0.104	0.132	0.091	0



Marmara Region, 1 each from the Central Anatolia, the Southeastern Region, and Italy and 3 genotypes from an unknown location. Although there was no clear discrimination in the phylogeny analysis based on location, this result was consistent with the model-based population structure.

Population structure analysis using ADMIXTURE software was performed to understand the genetic structure of 341 grapevine

genotypes. The cross-validation errors were examined under the models with $K = 1-14$. Genotypes with an admixture coefficient higher than 0.70 were assigned to a group, while those lower than 0.7 were assigned to the admixture group. ADMIXTURE simulation showed that the minimum value of cross-validation errors was 0.45 when $K=9$ (Figure 5), indicating that grapevine genotypes were classified into nine clusters (K1-K9) based on ADMIXTURE results (Figure 6). The 116 assigned genotypes (approximately 34% of 341 genotypes) were structured into nine clusters, while the other 225 genotypes (66% of 341 genotypes), exhibited membership values lower than 0.70 retained in the admixed genotype cluster. Although population structure analysis did not show any evidence of clustering according to their geographic origin, some inferred clusters (K2, K3, K4, K6, K7, and K9) only contained Turkish genotypes. Most of the genotypes that originated outside of Türkiye were categorized as admixture forms with varying membership levels shared among groups. In contrast, seven, three, and one genotype from Italy, France, Spain, and the USA were located in K8, K5, and K1, respectively. Detailed information about the individual coefficient of participation in the K clusters can be found in Supplementary Table S5.

To estimate the level of genetic differentiation, the ADMIXTURE software was used to calculate pairwise F_{st} values (Table 1). F_{st} values ranged from 0.077 to 0.255, with a mean of 0.15. The highest level of differentiation was observed between K1 and K3 ($F_{st}=0.255$). The lowest level of differentiation, F_{st} , was found between K2 and K9 ($F_{st}=0.077$). F_{st} values between groups showed that there was moderate divergence across all groups. In the study we also analyzed F_{st} values for individual SNP loci between the nine populations identified by ADMIXTURE analysis (Figure 7). Using a threshold of $F_{st} >0.5$, a total of 3091 divergent loci were detected (Table S6). The K2 vs K7 comparison had the lowest number (14) of divergent loci, while the highest (350) was observed in K1 vs K7. Among these, 1889 SNPs (61%) were found associated with genes. Forty-six highly divergent loci (F_{ST} values = ~ 0.93) identified between K3 and K7 and thirty-four of them were associated with a gene which is related important metabolic processes such as LRR and NB-ARC domains-containing disease resistance protein, heavy metal transport/detoxification superfamily

TABLE 2 Analysis of molecular variance (AMOVA) for 341 grapevine genotypes by grouping them based on STRUCTURE software and their geographic origins.

Structure-based AMOVA					
	df	Sum Sq	Mean Sq	Var	%
Among populations	9	909560.2	101062.24	584.84	0.65
Within populations	331	29428197.3	88906.94	88906.94	99.35
Total	340	30337757.4	89228.7	89491.78	100
Origin-based AMOVA					
	df	Sum Sq	Mean Sq	Var	%
Among populations	13	1282965	98689.60	584.84	0.63
Within populations	327	29054793	88852.58	88906.94	99.37
Total	340	30337757	89228.70	89491.78	100

df, degrees of freedom; Sum Sq, Sum of squares; Mean Sq, Mean square; Var, Estimated variance.

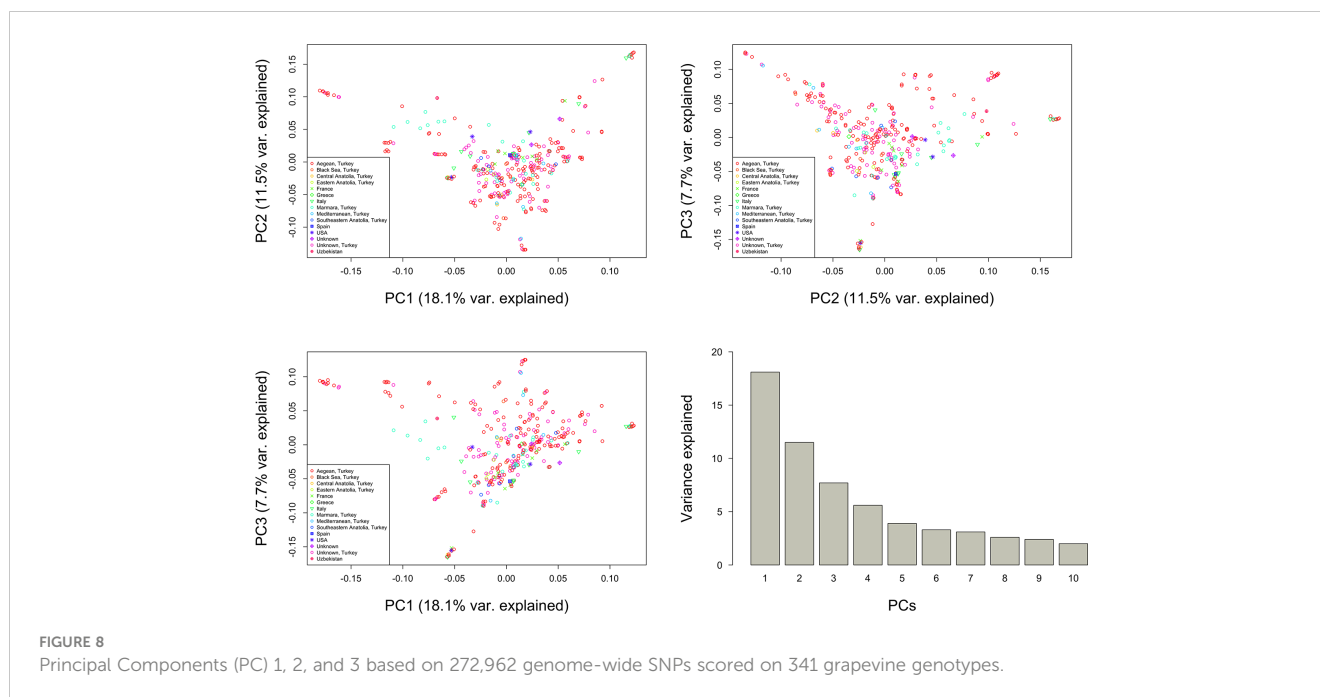


FIGURE 8

Principal Components (PC) 1, 2, and 3 based on 272,962 genome-wide SNPs scored on 341 grapevine genotypes.

protein, GATA type zinc finger transcription factor family protein, and AGD2-like defense response protein 1 (Table S6).

To understand the sources of genetic differentiation within and among the populations in the germplasm collection, we conducted an AMOVA based on the defined populations by structure analysis and geographic origin of the genotypes. The structure-based AMOVA results indicated that most of the observed genetic variation could be attributed to differences within populations (99.37%) rather than to the variation among populations (0.63%). Similarly, origin-based AMOVA also showed that the between population component of genetic variance is 0.65% and within population is 99.35% (Table 2).

The first three principal components (PCs) accounted for 37% of the genetic variation observed in the data (Figure 8). The first PC accounted for 18.1% of the observed variation, while the second and third PCs contributed 11.5% and 7.7%, respectively. Most genotypes

from different origins showed no clear separation across germplasm collection (Figure 8).

Linkage disequilibrium

LD was estimated for each pairwise SNP combination in the entire germplasm collection. A total of 2,481,581 (20.1%) pairs of markers showed a significant LD value (r^2) at $p < 0.01$, while 1,901,084 (15.4%) pairs of markers showed a significant LD at $P < 0.001$. The r^2 values for all significant loci ranged from 0.019 to 1, averaging 0.29. Only 131,151 (1.06%) pairs of markers showed complete LD ($r^2 = 1$). LD decay varied across the 19 chromosomes, as shown in Figure 9. LD in chromosome 16 decayed the most rapidly and slowly in chromosome 17. In general, LD decayed very fast when r^2 was between 0.45 and 0.2 and became flat when r^2 was 0.05 (Figure 9). The average LD decay for the entire genome was 30 kb when $r^2 = 0.2$. LD decay for a predicted r^2 of 0.2 varied from 20 to 60 kb based on the chromosomes.

Average pairwise r^2 values were similar among different chromosomes, ranging from 0.273 (chromosome 10) to 0.310 (chromosome 17) (Supplementary Table S7). Therefore, a number of significant and non-significant LD pairs and average r^2 values for them were given for each chromosome in Supplementary Table S7.

Discussion

The key requirements for progress in sustainable grapevine breeding programs depend upon the exploitation of genetic resources and characterization of genetic diversity as well as the investigation of genotype-phenotype associations by genomic and molecular techniques. For this purpose, GBS technique provides a

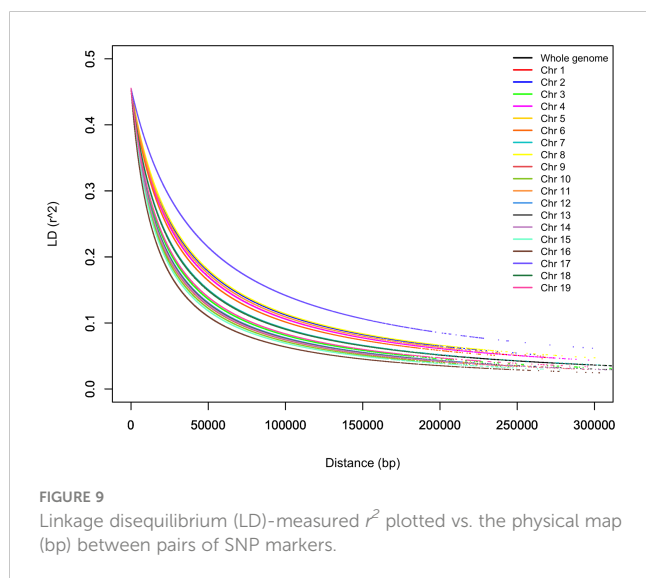


FIGURE 9

Linkage disequilibrium (LD)-measured r^2 plotted vs. the physical map (bp) between pairs of SNP markers.

large number of SNP markers distributed throughout the genome to enhance our knowledge of the genetic structure and genetic variability within grapevine germplasm collections.

The main objective of this study was to explore the genetic diversity and population structure of 341 grapevine genotypes from grapevine germplasm collection at Manisa Viticulture Research Institute using GBS-generated SNP markers. To our knowledge, no studies have been performed yet on Turkish grapevine germplasm based on high-throughput SNP discovery. This study is also the first comprehensive report in which the genetic diversity, population structure, and characterization of Turkish genotypes.

Genotyping by sequencing

High-throughput and cost-effective SNP discovery approaches such as GBS have become the most popular genotyping technique in plant species. They can be effectively used for various purposes in plant breeding and genetics (Elshire et al., 2011). In this study, GBS protocol enabled the discovery of thousands of SNPs to discriminate the 341 grapevine genotypes by providing many advantages, including relatively low cost, quick sample handling, and reduced genome complexity (Elshire et al., 2011; Barba et al., 2014). So far, several studies have used the GBS in domesticated and wild grapevine to assess genetic diversity (Klein et al., 2018; Miazzi et al., 2019; Miazzi et al., 2020), clarify evolutionary relationships (Klein et al., 2018), and identify marker-trait associations by QTL (Barba et al., 2014) and genome-wide association mapping (Guo et al., 2019).

Our study used one-enzyme (*ApeKI*) based GBS to genotype. *ApeKI* has comparatively few recognition sites in the major classes of plant retrotransposons, which helps reduce genome complexity (Elshire et al., 2011). Similarly, various studies preferred *ApeKI* enzyme to perform GBS in grapevine (Barba et al., 2014; Klein et al., 2018; Guo et al., 2019; Tello et al., 2019b). The average number of 3.1 million sequence reads per sample obtained in this study was much higher than other GBS studies in grapevine (Barba et al., 2014; Klein et al., 2018; Tello et al., 2019b). Klein et al. (Klein et al., 2018) reported 1.7 million sequence reads per sample by GBS of 359 *Vitis* genotypes. Averages of 712 thousand sequence reads per sample, and 2.8 million sequence reads per sample were obtained from the Illumina GA3 and HiSeq systems, respectively, by Barba et al. (Barba et al., 2014). Guo et al. (Guo et al., 2019) reported the highest number of sequence reads per sample, 4.8 million, in grapevine genotypes. These differences obtained from various studies were mainly due to different grapevine collections and the sequencing platform used for GBS. The average number of sequence reads per sample that was obtained here was also similar to what has been reported for various tree species: 3.5 million for apricot (Gürçan et al., 2016), 2.6 million for olive (Kaya et al., 2019), and 2.4 million for peach (Bielenberg et al., 2015).

After filtering based on missingness and MAF, a total of 272K high-quality SNPs were obtained, which is higher than the number of SNPs detected in other GBS studies in the grapevine (Barba et al.,

2014; Klein et al., 2018; Guo et al., 2019; Miazzi et al., 2020). Previously published studies in grapevine have produced a varying number of SNPs after filtering, such as 32K SNPs in 179 cultivars (Guo et al., 2019), 17K SNPs in 85 hybrid genotypes (Barba et al., 2014), 11K SNPs in 359 genotypes (Klein et al., 2018), and 25K SNPs in 40 genotypes (Miazzi et al., 2020) by GBS. In this study, high-quality SNPs covered all 19 chromosomes, and the number of SNPs per chromosome correlated highly with the chromosome size. A similar pattern of SNP distribution was reported by Guo et al. (Guo et al., 2019).

PIC is affected by marker types in addition to genetic diversity, genotype size, and breeding type of the species (Singh et al., 2013). While the PIC value of SNPs ranges from 0 to 0.5 due to the bi-allelic nature of the marker, it can be as high as 0.5–1.0 for multiallelic SSR markers (Singh et al., 2013; Chen et al., 2017). In this study, PIC values ranged from 0.0087 to 0.38, with an average value of 0.23. PIC value in a similar range was reported for SNP markers in a collection of grapevine genotypes by Cabezas et al. (Cabezas et al., 2011), Mihaljevic et al. (Žulj Mihaljević et al., 2020), and Lijavetzky et al. (Lijavetzky et al., 2007). Higher PIC values were reported in other grapevine studies with SSR markers (De Oliveira et al., 2020; Miazzi et al., 2020). Although the low PIC value of SNP markers is a limitation, it can be overcome by using a large number of SNPs (Lijavetzky et al., 2007). Although PIC values of SNP markers indicated low to moderate discriminating power in the study, 272K high-quality SNPs successfully assessed genetic diversity and investigated of genetic relatedness of the grapevine genotypes.

Observed heterozygosity (H_o) of the SNPs across 341 grapevine genotypes varied from 0.00 to 0.99 with an average of 0.23, which was slightly lower than the expected heterozygosity (H_e) that ranged from 0.08 to 0.50, with an average of 0.28, confirming the high level of heterozygosity in grapevine. The H_o range observed in this study was more comprehensive in range than 0.48 to 0.88 reported by Oliveira et al. (De Oliveira et al., 2020), 0.364 to 0.755 by Zdunic et al. (Zdunić et al., 2020), 0.32 to 0.94 by Mihaljevic et al. (Žulj Mihaljević et al., 2020). Mihaljevic et al. (Žulj Mihaljević et al., 2020) investigated the genetic diversity of the grapevine genotypes grown in Croatia and showed a mean observed heterozygosity of 0.71, greater than the mean expected 0.23 in this study. Furthermore, higher overall mean H_o values were obtained among cultivated grapevines in Portuguese (H_o = 0.361) (Cunha et al., 2020), Tunisia (H_o =0.325) (Ghaffari et al., 2014), Italy (H_o =0.791) (D'Onofrio, 2020), and various European countries from Central Europe and Part of the Western Balkan Peninsula (H_o = 0.578) (Zdunić et al., 2020). The H_e value obtained in this study was comparable with 0.304 and 0.348 reported by Ghaffari et al. (Ghaffari et al., 2014) and Cunha et al. (Cunha et al., 2020), respectively, while it was lesser than 0.698, 0.7, 0.777 reported by Zdunic et al. (Zdunić et al., 2020), Mihaljevic et al. (Žulj Mihaljević et al., 2020) and D'Onofrio (D'Onofrio, 2020). In line with our study, the expected heterozygosity value obtained from biallelic SNP markers (Ghaffari et al., 2014; De Lorenzis et al., 2019; Cunha et al., 2020) was lower than that of the studies which used

multiallelic markers such as SSRs (Augusto et al., 2021; Péros et al., 2021; Cretazzo et al., 2022).

The assessment of genetic diversity and population structure of the grapevine genotypes

For efficient conservation and utilization of germplasm collections, genetic diversity and population structure characterization have great significance (Ramanatha Rao and Hodgkin, 2002). Previous studies have been conducted on grapevine genetic characterization in the most important grapevine germplasm collections (Maul and Töpfer, 2015; Žulj Mihaljević et al., 2020). In this study, we have attempted to characterize 341 genotypes from grapevine germplasm collection in Manisa Viticulture Research Institute using GBS technology to enhance the conservation, development, and utilization of genetic resources in Türkiye.

In recent years, Viticulture Research Institutes in Türkiye significantly contributed to the conservation and use of Turkish grapevine germplasm collections. Ampelographic descriptors have characterized some important genotypes (Ates et al., 2011; Atak et al., 2014; Kara et al., 2018). Despite its high importance and implemented conservation strategies in Viticulture Research Institutes, molecular characterization of Turkish genotypes lagged behind other grape-producing countries. In the past, various grapevines, including both cultivated and wild genotypes from different geographical regions of Türkiye, have been analyzed using SSR markers to estimate the genetic diversity and structure (Vouillamoz et al., 2006; Selli et al., 2007; Tangolar et al., 2009; Boz et al., 2011; Yılmaz et al., 2020). A research project aimed to characterize 1150 Turkish genotypes in Tekirdağ Viticulture Research Institute by SSR markers; however, its results are not published (Gökbayrak and Söylemezoglu, 2010).

This is the first study in which the genetic diversity, structure, and characterization of 341 genotypes from Turkish germplasm collection, including national and international varieties such as Merlot, Cabernet Sauvignon, and Cardinal, were compared using GBS markers. Analyzed genotypes have agronomic and economic importance for the viticulture industry worldwide.

This study used complementary approaches to explore genetic diversity and population structure among grapevine genotypes. IBS-based genetic distance matrix and marker-based kinship coefficients were obtained to analyze genetic diversity using the three complementary clustering methods: Neighbor Joining (NJ)-based hierarchical clustering, Principal Coordinate Analysis (PCA), and Bayesian model-based clustering. In our study, the genetic distance coefficient ranged from 0.138 to 0.325, with an average dissimilarity of 0.277. The results showed that the 341 grapevine genotypes possessed a high genetic variation. Kinship values also demonstrate high genetic variation among most grapevine genotypes studied. The genetic dissimilarity values obtained in this study are comparable to those in other studies to characterize

of Turkish grapevine genotypes. Yılmaz et al. (Yılmaz et al., 2020) investigated 88 grapevine genotypes from the Central Anatolia region in Türkiye using SSR markers and reported genetic distances ranging from 0.056 to 0.207. Different genetic similarity/dissimilarity values among Turkish genotypes have been reported from different groups (Tangolar et al., 2009; Boz et al., 2011), which can be explained by the smaller size of the genotypes and different types of markers used in comparison with our study.

In this study, the neighbor-joining-based cluster analysis was implemented without considering the origin of the grapevine genotypes. Although grapevine genotypes were classified into seven major clusters and some genotypes were clustered as sub-clusters, the clustering pattern did not reflect the geographical origin of samples providing the common genetic background. Similar to our results, cluster analysis carried out in several other studies could not classify grapevine genotypes based on geographic origin (Ghaffari et al., 2014; De Lorenzis et al., 2019; Miazzi et al., 2020). However, Lorenzis et al. (De Lorenzis et al., 2015) and Zdunic et al. (Zdunić et al., 2020) were able to differentiate wild and cultivated grapevine genotypes with cluster analysis. Furthermore, differentiation between Western and Eastern European cultivars in phylogenetic trees was also shown by Zdunic et al. (Zdunić et al., 2020).

The PCA and structure analysis did not distinguish between the grapevine genotypes of different origins, highlighting the occurrence of gene flow and a high amount of admixture. Our results agree with Cipriani et al. (2010) for a collection of 1005 grapevine genotypes, reporting a weak correlation with their geographical origin based on the PCA. Several studies reported that PCA could separate wild and cultivated genotypes from different regions. Zdunić et al. (2020) conducted PCA based on 20 SSRs among 243 genotypes sampled in several European countries and obtained two major groups corresponding to the wild and cultivated grapevines. In another study, while PCA could distinguish 57 Tunisian grapevines into wild and cultivated genotypes, the groups did not seem to represent the geographic origins of the genotypes (Ghaffari et al., 2014).

In this study, 3091 divergent loci (F_{st} value > 0.5), including 49 highly divergent SNP loci (F_{st} value > 0.9), were identified by the estimation of pairwise F_{ST} values. The genes with the highest F_{st} value are involved in a wide variety of cellular processes, including transportation and detoxification of heavy metal ions (VIT_201s0011g04710), defence response of plants against infection (VIT_212s0055g00920) and biosynthesis of secondary metabolites and the regulation of gene expression (VIT_206s0004g07350) (Table S6). We identified the “VviERF045” gene (VIT_203s0063g00560, F_{st} value = 0.93), which belongs to the Ethylene Responsive Factor (ERF) family of transcription factors. VviERF045 is involved in the regulation of fruit ripening and response to biotic and abiotic stresses in grapevine {Leida, 2016, Insights into the role of the berry-specific ethylene responsive factor VviERF045}. Divergent loci associated with VIT_203s0063g00560, which encodes a protein, were annotated as a cadmium/zinc-

transporting ATPase in regulating cadmium uptake and transport in grapevine. Previously, Miazzi et al. {Miazzi et al., 2020, Marginal grapevine germplasm from Apulia (Southern Italy) represents an unexplored source of genetic diversity} identified divergent loci associated with genes involved in nitrogen metabolism, plant development, defense response, ripening process, stomatal movement, and carbohydrate metabolic processes. Similarly, Marrano et al. {Marrano et al., 2018, Genomic signatures of different adaptations to environmental stimuli between wild and cultivated *Vitis vinifera* L} also identified divergent loci associated with genes involved in the same metabolic pathway reported by Miazzi et al. {Miazzi et al., 2020, Marginal grapevine germplasm from Apulia (Southern Italy) represents an unexplored source of genetic diversity} In addition, several divergent loci, which were significantly differentiated between *sativa* and *sylvestris* populations found to be associated with genes involved in the adaptation to environmental changes {Marrano et al., 2018, Genomic signatures of different adaptations to environmental stimuli between wild and cultivated *Vitis vinifera* L}.

Our study found that AMOVA results indicated a high level of genetic differentiation within populations, while variation among populations was extremely low. This suggests that the high degree of differentiation within the population may be due to high levels of genetic exchange or gene flow. These findings are consistent with previous studies on grapevines that found high levels of diversity within populations but low levels of genetic diversity among populations (Ergül et al., 2011; Imazio et al., 2013; Riaz et al., 2018).

Since somatic mutations have more chance to happen and accumulate throughout the lifespan of perennial fruit crops, they create an important source of genetic variation (Torregrosa et al., 2011). Vegetative propagation is used to stabilize and propagate the grapevine somatic variants and vegetative descendant of a grapevine are called “clones” (Ibáñez et al., 2015). Although most clones of the same variety are identical, some can show phenotypic and genotypic differences over time (Roach et al., 2018). Phenotypic differences in the clones of the same cultivar have been reported in many studies (Roach et al., 2018; Grimplet et al., 2019; Vondras et al., 2019). Although clones are considered part of the same variety based on phenotypic characteristics, there is no clear distinction between clones and cultivars (Pelsy, 2010). Cases of synonymy (one genotype with several denominations) and homonymy (one denomination for several genotypes) can result from the clones of the same variety with phenotypic differences due to the somatic mutations and the mislabeling during the propagation of grapevines, respectively (Walker et al., 2006; De Oliveira et al., 2020). As in other grapevine-growing countries, the presence of homonyms and synonyms among local varieties has been reported in Turkish genotypes by ampelography and amperometry analysis (İşçi and Altındişli 2017; Arif et al., 2012) and molecular markers (Boz et al., 2011; Yılmaz et al., 2020). In our study, there are some genotypes with the same name that are adjacent to each other such as the two “Alarşın” (b-1-2 and c-8-8), the two “Ergin Çekirdeksizi” (a-31-4 and c-4-4), the two “Eşek Memesi” (c-7-2 and b-3-5), the two “Kara Dimrit” (a-23-1 and c-6-4), the two “Kınalı Ak Üzüm” (c-5-7 and b-2-5), the two “Silken Sarı” (a-19-6 and c-2-4). Although these genotypes were adjacent to each other, they had

an apparent genetic distance value suggesting that they are homonymous. A few genotypes are also very close to each other in the same cluster which has the same name, including the two “Benli Belercen” (c-2-6 and a-12-1), the two “Beyrut Hurması” (b-1-6 and c-3-2), the two “Çekirdeksiz Babası” (c-5-5 and b-1-8), the two “Hasan Üzümü” (c-9-4 and b-2-2), the two “Ufak Dimrit” (a-17-6 and b-4-7). Although the two “Paşalar Üzümü” (a-1-2 and c-7-4) are in the same cluster, they are not close to each other. Many genotypes, such as the two “Foça Razakısı” (c-7-5 and b-4-2), the two “İnek Memesi” (b-3-4 and c-4-3), the two “Nezi Üzümü” (b-2-8 and c-10-5), the two “Şam Üzümü” (b-3-7 and c-11-5) and the two Regina (b-6-6 and c-3-6) have same names but different origins and did not cluster in the dendrogram and seemed to be homonymous. Possible synonyms and homonyms in Turkish grapevine germplasm have been previously reported by Ergül et al. (Ergül et al., 2006), Vouillamoz et al. (Vouillamoz et al., 2006), Tangolar et al. (Tangolar et al., 2009), Boz et al. (Boz et al., 2011), İşçi et al. (Burçak and DİLLİ, 2015) and Yılmaz et al. (Yılmaz et al., 2020). Genotyping in these studies is mainly based on SSR markers which are very useful for genetic diversity evaluation and cultivar identification, including the detection of synonyms and homonyms (Cipriani et al., 2010; Laucou et al., 2011; Emanuelli et al., 2013). The genetic diversity among the central Anatolian grapevine genotypes was recently characterized using SSRs by Yılmaz et al. (Yılmaz et al., 2020). Their study reported two cases of identical, seven cases of homonymous, and nine cases of synonymous grape genotypes based on 17 SSR markers (Yılmaz et al., 2020). In another study, one synonym and four homonymies were identified among 55 grape cultivars originating from the Southeastern Region using 14 SSR markers (Boz et al., 2011). Although SSR markers are highly informative, reproducible, codominant, and multi-allelic, their limitations in the accurate identification of clones have been also reported (González-Techera et al., 2004; Pelsy et al., 2010; Mercati et al., 2016). Homonym and synonym cultivars were also reported in Spanish (Moreno-Sanz et al., 2008), Iranian (Fatahi et al., 2003), Armenian (Margaryan et al., 2021), and Italian (Cipriani et al., 2010) cultivars.

Linkage disequilibrium

LD is the non-random association between alleles at different sites and depends on many factors such as recombination, mutation, genetic drift, selection, and population admixture and size (Flint-Garcia et al., 2003). While LD decays more rapidly in cross-pollinated plants, self-pollinated plants show less decay of LD related to less effective recombination (Campoy et al., 2016). In this study, a comprehensive genome-wide LD in grapevine was evaluated by high-density SNPs (272K) over 341 genotypes. r^2 values decreased to 0.2 within 30 kb and 0.1 within 90 kb for the whole genome. This result is expected and consistent with the trend of rapid LD decay in grapevines (Myles et al., 2011; Nicolas et al., 2016; Guo et al., 2019). The First LD study in grapevine was reported by Barnaud et al. (Barnaud et al., 2006) using SSR markers. Up to 16.8 cM large extent of LD was observed within linkage groups. Myles et al. (Myles et al., 2011) observed that the decay of LD was down to 0.2 within less than

10 kb using the Vitis9K SNP array. Lijavetzky et al. (Lijavetzky et al., 2007) evaluated the LD decay with the 1500 SNPs and observed a lower level of LD which r^2 values decreased to 0.2 within 250 bp. In another study, LD reached 0.2 within 10kb using 26K SNPs (Marrano et al., 2018). Guo et al. (Guo et al., 2019) observed faster LD decay which r^2 reached 0.03 within 16.6 kb using 32K SNPs by GBS. In our study, significant variation in LD decay was found among chromosomes; the decay of LD was down to 0.2 ranged from 20 to 60 kb. Nicolas et al. (Nicolas et al., 2016) obtained variability in LD decay changing based on the subgroups and genomic regions. LD decay reached 0.2 and varied from 9 to 458 kb (Nicolas et al., 2016).

To assess LD in wild *V. vinifera*, 85 plants from Southern France were analyzed by 36 SSR markers, and LD was slower in cultivated than in wild grapevine (Barnaud et al., 2010). It was reported that this difference between cultivated and wild grapevine is derived from the vegetative propagation and domestication bottlenecks (Barnaud et al., 2010). By contrast, Myles et al. (Myles et al., 2011) reported a consistent trend of rapid LD decay between wild and domesticated grapevine. Marrano et al. (Marrano et al., 2017) also analyzed wild and domesticated grapevines using 14K high-quality SNPs by RAD sequencing. Vitis20K SNP array was also used for genotyping of the same grapevines by Marrano et al. (Marrano et al., 2018), and they combined SNPs (26K) from both RAD sequencing and the Vitis20K array for LD analysis. In contrast to Barnaud et al. (Barnaud et al., 2010) and Myles et al. (Myles et al., 2011), slower LD decay was observed within wild grapevines, which r^2 reached values below 0.2 within 20kb, LD (r^2) decayed below 0.2 within 10 kb in domesticated grapevines based on RAD sequencing. When they used combined data (26K), the decay of LD appeared slower within the wild group, where r^2 reached values below 0.2 within 20 kb (Marrano et al., 2018).

These discrepancies could be attributed to different types or varying numbers of markers that affect the genome's coverage (Marrano et al., 2018). In addition, since different factors such as mutation, recombination, selection, mating system, and population contribute to LD, variation in LD estimates may also be related to these factors (Flint-Garcia et al., 2003).

Conclusion

This study used high-throughput GBS technology to explore genetic diversity, population structure among grapevine genotypes, and the possibility of using SNP markers for genetic analyses in genetic enhancement. The hierarchical cluster and principal components analysis indicated a mixing of grapevines from different regions, suggesting the strong gene flow among genotypes. This is in line with the results of population structure analysis which showed a high amount of genetic admixture in Turkish genotypes. Information on genetic diversity and population structure is essential to provide valuable genetic resources for grapevine genetic improvement and protect and develop this vital crop from erosion. Furthermore, extensive genetic diversity should be identified and maintained for a sustainable viticulture industry to develop new grape cultivars with desired traits through marker-assisted breeding. Identifying divergent loci could provide useful

information for developing grapevine varieties that adapt effectively to new environmental conditions.

Data availability statement

All the raw sequencing reads for all genotypes have been submitted to the NCBI Sequence Read Archive and deposited under the "BioProject ID": PRJNA742054.

Author contributions

HK, YD, and AÜ conceived and designed the study. HK, YD, and TO performed the experiments. HK analyzed the data and prepared figures and tables. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Manisa Celal Bayar University Scientific Research Projects funds (2017-113 to HK and TO) and (TAGEMTBAD-16-A01-P01-012 to YD, and AÜ).

Acknowledgments

The authors thank M. Temel, S. Pamay, and A. Abacı for technical assistance. We thank Dr. Deniz Akdemir and Dr. Roberto Lozano for helpful suggestions on statistical analysis. Special thanks to Bilgehan Nevruz for technical assistance on statistical analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1121811/full#supplementary-material>

References

- Adam-Blondon, A.-F., Jaillon, O., Vezzulli, S., Zharkikh, A., Troglio, M., Velasco, R., et al. (2011). Genome sequence initiatives. *Genetics Genomics Breed. Grapes* 211, 234.
- Adam-Blondon, A.-F., Roux, C., Claux, D., Butterlin, G., Merdinoglu, D., and This, P. (2004). Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. *Theor. Appl. Genet.* 109 (5), 1017–1027. doi: 10.1007/s00122-004-1704-y
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19 (9), 1655–1664. doi: 10.1101/gr.094052.109
- Ali, K., Maltese, F., Choi, Y. H., and Verpoorte, R. (2010). Metabolic constituents of grapevine and grape-derived products. *Phytochem. Rev.* 9 (3), 357–378. doi: 10.1007/s11101-009-9158-0
- Arif, A., Ahmet, A., Ali, F. G., and Cengiz, O. (2012). Molecular and ampelographic characterization of some grape hybrids (*Vitis vinifera* L.). *Afr. J. Agric. Res.* 7 (33), 4596–4606. doi: 10.5897/AJAR11.1957
- Atak, A., Kahraman, K., and Söylemezoğlu, G. (2014). Ampelographic identification and comparison of some table grape (*Vitis vinifera* L.) clones. *New Z. J. Crop Hortic. Sci.* 42 (2), 77–86. doi: 10.1080/01140671.2013.851092
- Ates, F., Coban, H., Kara, Z., and Sabir, A. (2011). Ampelographic characterization of some grape cultivars (*Vitis vinifera* L.) grown in south-western region of Turkey. *Bulg. J. Agric. Sci.* 17 (3), 314–324.
- Augusto, D., Ibáñez, J., Pinto-Sintra, A. L., Falco, V., Leal, F., Martínez-Zapater, J. M., et al. (2021). Grapevine diversity and genetic relationships in northeast Portugal old vineyards. *Plants* 10 (12), 2755. doi: 10.3390/plants10122755
- Barba, P., Cadle-Davidson, L., Harriman, J., Glaubitz, J. C., Brooks, S., Hyma, K., et al. (2014). Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. *Theor. Appl. Genet.* 127 (1), 73–84. doi: 10.1007/s00122-013-2202-x
- Barnaud, A., Lacombe, T., and Doligez, A. (2006). Linkage disequilibrium in cultivated grapevine, *Vitis vinifera* L. *Theor. Appl. Genet.* 112 (4), 708–716. doi: 10.1007/s00122-005-0174-1
- Barnaud, A., Laucou, V., This, P., Lacombe, T., and Doligez, A. (2010). Linkage disequilibrium in wild French grapevine, *Vitis vinifera* L. subsp. *silvestris*. *Heredity* 104 (5), 431–437. doi: 10.1038/hdy.2009.143
- Bielenberg, D. G., Rauh, B., Fan, S., Gasic, K., Abbott, A. G., Reighard, G. L., et al. (2015). Genotyping by sequencing for SNP-based linkage map construction and QTL analysis of chilling requirement and bloom date in peach [*Prunus persica* (L.) batsch]. *PLoS One* 10 (10), e0139406. doi: 10.1371/journal.pone.0139406
- Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32 (3), 314.
- Bouby, L., Figueiral, I., Bouchette, A., Rovira, N., Ivorra, S., Lacombe, T., et al. (2013). Bioarchaeological insights into the process of domestication of grapevine (*Vitis vinifera* L.) during Roman times in southern France. *PLoS One* 8 (5), e63195. doi: 10.1371/journal.pone.0063195
- Boz, Y., Bakir, M., Çelikkol, B., Kazan, K., Yilmaz, F., Cakir, B., et al. (2011). Genetic characterization of grape (*Vitis vinifera* L.) germplasm from southeast Anatolia by SSR markers. *Vitis* 50 (3), 99–106.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23 (19), 2633–2635. doi: 10.1093/bioinformatics/btm308
- Browning, B. L., and Browning, S. R. (2016). Genotype imputation with millions of reference samples. *Am. J. Hum. Genet.* 98 (1), 116–126. doi: 10.1016/j.ajhg.2015.11.020
- Burçak, İ., and DİLLİ, Y. (2015). Characterization of autochthonous grapevine cultivars (*Vitis vinifera* L.) from the Aegean region of Turkey using simple sequence repeats (SSRs). *J. Agric. Sci.* 21 (4), 538–545. doi: 10.1501/Tarimbil_0000001355
- Cabezas, J. A., Ibáñez, J., Lijavetzky, D., Vélez, D., Bravo, G., Rodriguez, V., et al. (2011). A 48 SNP set for grapevine cultivar identification. *BMC Plant Biol.* 11 (1), 1–12. doi: 10.1186/1471-2229-11-153
- Campoy, J. A., Lerigoleur-Balsemin, E., Christmann, H., Beauvieux, R., Girollet, N., Quero-García, J., et al. (2016). Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. *BMC Plant Biol.* 16 (1), 1–15. doi: 10.1186/s12870-016-0712-9
- Chen, W., Hou, L., Zhang, Z., Pang, X., and Li, Y. (2017). Genetic diversity, population structure, and linkage disequilibrium of a core collection of *Ziziphus jujuba* assessed with genome-wide SNPs developed by genotyping-by-sequencing and SSR markers. *Front. Plant Sci.* 8, 575. doi: 10.3389/fpls.2017.00575
- Cipriani, G., Spadotto, A., Jurman, I., Di Gaspero, G., Crespan, M., Meneghetti, S., et al. (2010). The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. *Theor. Appl. Genet.* 121 (8), 1569–1585. doi: 10.1007/s00122-010-1411-9
- Cretazzo, E., Moreno Sanz, P., Lorenzi, S., Benítez, M. L., Velasco, L., and Emanuelli, F. (2022). Genetic characterization by SSR markers of a comprehensive wine grape collection conserved at rancho de la merced (Andalusia, Spain). *Plants* 11 (8), 1088. doi: 10.3390/plants11081088
- Crossa, J., Beyene, Y., Kassa, S., Pérez, P., Hickey, J. M., Chen, C., et al. (2013). Genomic prediction in maize breeding populations with genotyping-by-sequencing. *G3: Genes Genomes Genet.* 3 (11), 1903–1926. doi: 10.1534/g3.113.008227
- Cunha, J., Ibáñez, J., Teixeira-Santos, M., Brazão, J., Feveiro, P., Martínez-Zapater, J. M., et al. (2020). Genetic relationships among Portuguese cultivated and wild *Vitis vinifera* L. germplasm. *Front. Plant Sci.* 11, 127. doi: 10.3389/fpls.2020.00127
- D'Agostino, N., Taranto, F., Camposo, S., Mangini, G., Fanelli, V., Gadaleta, S., et al. (2018). GBS-derived SNP catalogue unveiled wide genetic variability and geographical relationships of Italian olive cultivars. *Sci. Rep.* 8 (1), 15877. doi: 10.1038/s41598-018-34207-y
- Dalbó, M., Ye, G., Weeden, N., Steinkellner, H., Sefc, K., and Reisch, B. (2000). A gene controlling sex in grapevines placed on a molecular marker-based genetic map. *Genome* 43 (2), 333–340. doi: 10.1139/g99-136
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27 (15), 2156–2158. doi: 10.1093/bioinformatics/btr330
- De Lorenzis, G., Chipashvili, R., Failla, O., and Maghradze, D. (2015). Study of genetic variability in *Vitis vinifera* L. germplasm by high-throughput Vitis18kSNP array: the case of Georgian genetic resources. *BMC Plant Biol.* 15 (1), 1–14. doi: 10.1186/s12870-015-0510-9
- De Lorenzis, G., Mercati, F., Bergamini, C., Cardone, M. F., Lupini, A., Mauceri, A., et al. (2019). SNP genotyping elucidates the genetic diversity of magna graecia grapevine germplasm and its historical origin and dissemination. *BMC Plant Biol.* 19 (1), 1–15. doi: 10.1186/s12870-018-1576-y
- De Oliveira, G. L., De Souza, A. P., De Oliveira, F. A., Zucchi, M. I., De Souza, L. M., and Moura, M. F. (2020). Genetic structure and molecular diversity of Brazilian grapevine germplasm: management and use in breeding programs. *PLoS One* 15 (10), e0240665. doi: 10.1371/journal.pone.0240665
- Deschamps, S., Llaça, V., and May, G. D. (2012). Genotyping-by-sequencing in plants. *Biology* 1 (3), 460–483. doi: 10.3390/biology1030460
- Di Gaspero, G., and Cattonaro, F. (2010). Application of genomics to grapevine improvement. *Aust. J. Grape Wine Res.* 16, 122–130. doi: 10.1111/j.1755-0238.2009.00072.x
- D'Onofrio, C. (2020). Introgression among cultivated and wild grapevine in Tuscany. *Front. Plant Sci.* 11, 202. doi: 10.3389/fpls.2020.00202
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6 (5), e19379. doi: 10.1371/journal.pone.0019379
- Emanuelli, F., Lorenzi, S., Grzeskowiak, L., Catalano, V., Stefanini, M., Troglio, M., et al. (2013). Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. *BMC Plant Biol.* 13 (1), 1–17. doi: 10.1186/1471-2229-13-39
- Ergül, A., Kazan, K., Aras, S., Çevik, V., Çelik, H., and Söylemezoğlu, G. (2006). AFLP analysis of genetic variation within the two economically important Anatolian grapevine (*Vitis vinifera* L.) varietal groups. *Genome* 49 (5), 467–475. doi: 10.1139/g05-121
- Ergül, A., Perez-Rivera, G., Söylemezoğlu, G., Kazan, K., and Arroyo-García, R. (2011). Genetic diversity in Anatolian wild grapes (*Vitis vinifera* subsp. *silvestris*) estimated by SSR markers. *Plant Genet. Resour.* 9 (3), 375–383. doi: 10.1017/S1479262111000013
- FAOSTAT. (2018). Statistics of the Food and Agriculture Organization of the United Nations. In: Food Agric. Organ. United Nations.
- Fatahi, R., Ebadi, A., Bassil, N., Mehlenbacher, S., and Zamani, Z. (2003). Characterization of Iranian grapevine cultivars using microsatellite markers. *VITIS-GEILWEILERHOF* 42 (4), 185–192.
- Flint-García, S. A., Thornsberry, J. M., and Buckler, IV, E.S. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54 (1), 357–374. doi: 10.1146/annurev.arplant.54.031902.134907
- Flutre, T., Le Cunff, L., Fodor, A., Launay, A., Romieu, C., Berger, G., et al. (2020). Genome-wide association and prediction studies using a grapevine diversity panel give insights into the genetic architecture of several traits of interest. *bioRxiv*. doi: 10.1101/2020.09.10.290890
- Folts, J. D. (2002). "Potential health benefits from the flavonoids in grape products on vascular disease." in *Adv. Exp. Med. Biol.* 505, 95–111. doi: 10.1007/978-1-4757-5235-9_9
- Gago, P., Boso, S., Santiago, J. L., Soler, J.-X., Peiró, R., García, J., et al. (2022). Characterization of grapevine genetic resources in the comunitat valenciana (Spain). *Int. J. Fruit Sci.* 22 (1), 287–302. doi: 10.1080/15538362.2022.2037039
- Gardner, K. M., Brown, P., Cooke, T. F., Cann, S., Costa, F., Bustamante, C., et al. (2014). Fast and cost-effective genetic mapping in apple using next-generation sequencing. *G3: Genes Genomes Genet.* 4 (9), 1681–1687. doi: 10.1534/g3.114.011023
- Ghaffari, S., Hasnaoui, N., Zinelabidine, L. H., Ferchichi, A., Martínez-Zapater, J. M., and Ibáñez, J. (2014). Genetic diversity and parentage of Tunisian wild and cultivated grapevines (*Vitis vinifera* L.) as revealed by single nucleotide polymorphism (SNP) markers. *Tree Genet. Genomes* 10 (4), 1103–1112. doi: 10.1007/s11295-014-0746-9

- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9 (2), e90346. doi: 10.1371/journal.pone.0090346
- Gökbayrak, Z., and Söylemezoglu, G. (2010). Grapevine throughout the history of Anatolia. *Int. J. Bot.* 6 (4), 465–472. doi: 10.3923/ijb.2010.465.472
- González-Tejera, A., Jubany, S., Ponce de Leon, I., Boido, E., Dellacassa, E., Carrau, F., et al. (2004). Molecular diversity within clones of cv. tannat (*Vitis vinifera*). *Vitis* 43 (4), 179–185.
- Goonetilleke, S. N., March, T. J., Wirthensohn, M. G., Arús, P., Walker, A. R., and Mather, D. E. (2018). Genotyping by sequencing in almond: SNP discovery, linkage mapping, and marker design. *G3: Genes Genomes Genet.* 8 (1), 161–172. doi: 10.1534/g3.117.300376
- Granato, I. S., Galli, G., de Oliveira Couto, E. G., Mendonça, L. F., and Fritsche-Neto, R. (2018). snpReady: a tool to assist breeders in genomic analysis. *Mol. Breed.* 38 (8), 1–7. doi: 10.1007/s11032-018-0844-8
- Grimplet, J., Ibáñez, S., Baroja, E., Tello, J., and Ibáñez, J. (2019). Phenotypic, hormonal, and genomic variation among vitis vinifera clones with different cluster compactness and reproductive performance. *Front. Plant Sci.* 9, 1917. doi: 10.3389/fpls.2018.01917
- Guajardo, V., Solís, S., Sagredo, B., Gainza, F., Muñoz, C., Gasic, K., et al. (2015). Construction of high density sweet cherry (*Prunus avium* L.) linkage maps using microsatellite markers and SNPs detected by genotyping-by-sequencing (GBS). *PLoS One* 10 (5), e0127750. doi: 10.1371/journal.pone.0127750
- Guo, D.-L., Zhao, H.-L., Li, Q., Zhang, G.-H., Jiang, J.-F., Liu, C.-H., et al. (2019). Genome-wide association study of berry-related traits in grape [*Vitis vinifera* L.] based on genotyping-by-sequencing markers. *Horticulture Res.* 6 (1), 1–13. doi: 10.1038/s41438-018-0089-z
- Gürkan, K., Teber, S., Ercisli, S., and Yilmaz, K. U. (2016). Genotyping by sequencing (GBS) in apricots and genetic diversity assessment with GBS-derived single-nucleotide polymorphisms (SNPs). *Biochem. Genet.* 54 (6), 854–885. doi: 10.1007/s10528-016-9762-9
- Hill, W., and Weir, B. (1988). Variances and covariances of squared linkage disequilibria in finite populations. *Theor. population Biol.* 33 (1), 54–78. doi: 10.1016/0040-5809(88)90004-4
- Ibáñez, J., Carreño, J., Yuste, J., and Martínez-Zapater, J. (2015). Grapevine breeding and clonal selection programmes in Spain. *Grapevine Breed. Programs for the Wine Industry*, 9 183–209. doi: 10.1016/B978-1-78242-075-0.00009-0
- Imazio, S., Maghradze, D., De Lorenzis, G., Bacilieri, R., Laucou, V., This, P., et al. (2013). From the cradle of grapevine domestication: molecular overview and description of Georgian grapevine (*Vitis vinifera* L.) germplasm. *Tree Genet. Genomes* 9 (3), 641–658. doi: 10.1007/s11295-013-0597-9
- İşçi, B., and Altundışli, A. (2017). Ampelographic characterization of Turkish indigenous grape accessions and European cultivars (*Vitis vinifera* L.). *Int. J. Agric. Environ. Food Sci.* 1 (1), 1–16. doi: 10.31015/ijaefs.17001
- Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449 (7161), 463. doi: 10.1038/nature06148
- Jiang, H., Lei, R., Ding, S.-W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinf.* 15 (1), 1–12. doi: 10.1186/1471-2105-15-182
- Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2, e281. doi: 10.7717/peerj.281
- Kara, Z., Sabir, A., and Eker, Ö. (2018). Ancient grape vitis vinifera L. cv 'Ekşi kara' in Anatolia. *Selçuk J. Agric. Food Sci.* 32 (3), 416–423. doi: 10.15316/SJAFS.2018.116
- Kaya, H. B., Akdemir, D., Lozano, R., Cetin, O., Kaya, H. S., Sahin, M., et al. (2019). Genome wide association study of 5 agronomic traits in olive (*Olea europaea* L.). *Sci. Rep.* 9 (1), 1–14. doi: 10.1038/s41598-019-55338-w
- Klein, L. L., Miller, A. J., Ciotir, C., Hyma, K., Uribe-Convers, S., and Londo, J. (2018). High-throughput sequencing data clarify evolutionary relationships among north American vitis species and improve identification in USDA vitis germplasm collections. *Am. J. Bot.* 105 (2), 215–226. doi: 10.1002/ajb2.1033
- Lacombe, T., Boursiquot, J.-M., Laucou, V., Dechesne, F., Varès, D., and This, P. (2007). Relationships and genetic diversity within the accessions related to malvasia held in the domaine de vassal grape germplasm repository. *Am. J. Enology Viticulture* 58 (1), 124–131. doi: 10.5344/ajev.2007.58.1.124
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9 (4), 357. doi: 10.1038/nmeth.1923
- Laucou, V., Lacombe, T., Dechesne, F., Siret, R., Bruno, J.-P., Dessup, M., et al. (2011). High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theor. Appl. Genet.* 122 (6), 1233–1245. doi: 10.1007/s00122-010-1527-y
- Laucou, V., Launay, A., Bacilieri, R., Lacombe, T., Adam-Blondon, A.-F., Bérard, A., et al. (2018). Extended diversity analysis of cultivated grapevine vitis vinifera with 10K genome-wide SNPs. *PLoS One* 13 (2), e0192540. doi: 10.1371/journal.pone.0192540
- Leida, C., Dal Rì, A., Dalla Costa, L., Gómez, M. D., Pompili, V., Sonogo, P., et al (2016). Insights into the role of the berry-specific ethylene responsive factor VvIERF045. *Front. Plant Sci.* 7, 1793.
- Letunic, I., and Bork, P. (2019). Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47 (W1), W256–W259. doi: 10.1093/nar/gkz239
- Lijavetzky, D., Cabezas, J. A., Ibáñez, A., Rodríguez, V., and Martínez-Zapater, J. M. (2007). High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *BMC Genomics* 8 (1), 1–11. doi: 10.1186/1471-2164-8-424
- Lodhi, M. A., Ye, G.-N., Weeden, N. F., and Reisch, B. I. (1994). A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. *Plant Mol. Biol. Rep.* 12 (1), 6–13. doi: 10.1007/BF02668658
- Lodhi, M., Ye, G.-N., Weeden, N., Reisch, B., and Daly, M. (1995). A molecular marker based linkage map of vitis. *Genome* 38 (4), 786–794. doi: 10.1139/g95-100
- Lopes, M. S., Sefc, K. M., Dias, E. E., Steinkellner, H., Machado, M. L. C., and Machado, A. C. (1999). The use of microsatellites for germplasm management in a Portuguese grapevine collection. *Theor. Appl. Genet.* 99 (3–4), 733–739. doi: 10.1007/s001220051291
- Margaryan, K., Melyan, G., Röckel, F., Töpfer, R., and Maul, E. (2021). Genetic diversity of Armenian grapevine (*Vitis vinifera* L.) germplasm: molecular characterization and parentage analysis. *Biology* 10 (12), 1279. doi: 10.3390/biology10121279
- Marrano, A., Birolo, G., Prazzoli, M. L., Lorenzi, S., Valle, G., and Grando, M. S. (2017). SNP-discovery by RAD-sequencing in a germplasm collection of wild and cultivated grapevines (*V. vinifera* L.). *PLoS One* 12 (1), e0170655. doi: 10.1371/journal.pone.0170655
- Marrano, A., Micheletti, D., Lorenzi, S., Neale, D., and Grando, M. S. (2018). Genomic signatures of different adaptations to environmental stimuli between wild and cultivated vitis vinifera L. *Horticulture Res.* 5(34). doi: 10.1038/s41438-018-0041-2
- Martin, J. P., Pinto-Carnide, O., Leal, F., Martínez, M., and Ortiz, J. M. (2006). Determination of relationships among autochthonous grapevine varieties (*Vitis vinifera* L.) in the Northwest of the Iberian peninsula by using microsatellite markers. *Genet. Resour. Crop Evol.* 53 (6), 1255–1261. doi: 10.1007/s10722-005-5679-6
- Maul, E., Sudharma, K. N., Kecke, S., Marx, G., Müller, C., Audeguin, L., et al. (2012). The European vitis database (www.eu-vitis.de): a technical innovation through an online uploading and interactive modification system. *Vitis* 51 (2), 79–85. doi: 10.1051/bioconf/20150501009
- Maul, E., and Töpfer, R. (2015). "Vitis international variety catalogue (VIVC): a cultivar database referenced by genetic profiles and morphology," in *BIO web of conferences* (EDP Sciences) 5, 01009. doi: 10.1051/bioconf/20150501009
- McGovern, P. E. (2003). The search for the origins of viticulture. *The Origins and Ancient History of Wine* P McGovern, S Fleming and S Katz (Princeton University Press) 19 (4), 619–628. doi: 10.1016/j.gpb.2020.10.007
- Mercati, F., De Lorenzis, G., Brancadoro, L., Lupini, A., Abenavoli, M. R., Barbagallo, M. G., et al. (2016). High-throughput 18K SNP array to assess genetic variability of the main grapevine cultivars from Sicily. *Tree Genet. Genomes* 12 (3), 1–15. doi: 10.1007/s11295-016-1021-z
- Miazzì, M., D'Agostino, N., Gadaleta, S., Di Rienzo, V., Fanelli, V., Sabetta, W., et al. (2019). Genotyping-by-sequencing-derived single-nucleotide polymorphism catalog from a grapevine (*Vitis vinifera* L.) germplasm collection that includes the most representative apulian autochthonous cultivars. *Acta Hort.* 1248. doi: 10.17660/ActaHortic.2019.1248.10
- Miazzì, M. M., D'Agostino, N., di Rienzo, V., Venerito, P., Savino, V. N., Fucilli, V., et al. (2020). Marginal grapevine germplasm from apulia (Southern Italy) represents an unexplored source of genetic diversity. *Agronomy* 10 (4), 563. doi: 10.3390/agronomy10040563
- Moreno-Sanz, P., Suárez, B., and Loureiro, M. (2008). Identification of synonyms and homonyms in grapevine cultivars (*Vitis vinifera* L.) from Asturias (Spain). *J. Hort. Sci. Biotechnol.* 83 (6), 683–688. doi: 10.1080/14620316.2008.11512444
- Myles, S., Boyko, A. R., Owens, C. L., Brown, P. J., Grassi, F., Aradhya, M. K., et al. (2011). Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci.* 108 (9), 3530–3535. doi: 10.1073/pnas.1009363108
- Nicolas, S. D., Péros, J.-P., Lacombe, T., Launay, A., Le Paslier, M.-C., Bérard, A., et al. (2016). Genetic diversity, linkage disequilibrium and power of a large grapevine (*Vitis vinifera* L.) diversity panel newly designed for association studies. *BMC Plant Biol.* 16 (1), 1–19. doi: 10.1186/s12870-016-0754-z
- Pelsy, F. (2010). Molecular and cellular mechanisms of diversity within grapevine varieties. *Heredity* 104 (4), 331–340. doi: 10.1038/hdy.2009.161
- Pelsy, F., Hocquigny, S., Moncada, X., Barbeau, G., Forget, D., Hinrichsen, P., et al. (2010). An extensive study of the genetic diversity within seven French wine grape variety collections. *Theor. Appl. Genet.* 120 (6), 1219–1231. doi: 10.1007/s00122-009-1250-8
- Péros, J. P., Cousins, P., Launay, A., Cubry, P., Walker, A., Prado, E., et al. (2021). Genetic diversity and population structure in vitis species illustrate phylogeographic patterns in eastern north America. *Mol. Ecol.* 30 (10), 2333–2348. doi: 10.1111/mec.15881

- Ramanatha Rao, V., and Hodgkin, T. (2002). Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell Tissue Organ Cult.* 68 (1), 1–19. doi: 10.1023/A:1013359015812
- Riaz, S., Dangl, G., Edwards, K., and Meredith, C. (2004). A microsatellite marker based framework linkage map of vitis vinifera l. *Theor. Appl. Genet.* 108 (5), 864–872. doi: 10.1007/s00122-003-1488-5
- Riaz, S., De Lorenzis, G., Velasco, D., Koehmstedt, A., Maghradze, D., Bobokashvili, Z., et al. (2018). Genetic diversity analysis of cultivated and wild grapevine (*Vitis vinifera* l.) accessions around the Mediterranean basin and central Asia. *BMC Plant Biol.* 18 (1), 1–14. doi: 10.1186/s12870-018-1351-0
- Roach, M. J., Johnson, D. L., Bohlmann, J., van Vuuren, H. J., Jones, S. J., Pretorius, I. S., et al. (2018). Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay. *PLoS Genet.* 14 (11), e1007807. doi: 10.1371/journal.pgen.1007807
- Selli, F., Bakir, M., Inan, G., Aygun, H., Boz, Y., Yasasin, A., et al. (2007). Simple sequence repeat-based assessment of genetic diversity in 'Dimrit' and 'Genre' grapevine accessions from Turkey. *Vitis* 46 (4), 182–187.
- Singh, N., Choudhury, D. R., Singh, A. K., Kumar, S., Srinivasan, K., Tyagi, R., et al. (2013). Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. *PLoS One* 8 (12), e84136. doi: 10.1371/journal.pone.0084136
- Söylemezoğlu, G., Atak, A., Boz, Y., Ünal, A., and Sağlam, M. (2016). Viticulture in Turkey. *Chronica Hort.* 56 (2), 27–31.
- Tangolar, S. G., Soydam, S., Bakir, M., Karaagac, E., Tangolar, S., and Ergul, A. (2009). Genetic analysis of grapevine cultivars from the eastern Mediterranean region of Turkey, based on SSR markers. *Tarim Bilim Derg* 15 (1), 1–8.
- Team, R.C. (2016). R: a language and environment for statistical computing. *R Found. Stat. Comput* 1, 409.
- Teh, C.-K., Ong, A.-L., Kwong, Q.-B., Apparow, S., Chew, F.-T., Mayes, S., et al. (2016). Genome-wide association study identifies three key loci for high mesocarp oil content in perennial crop oil palm. *Sci. Rep.* 6 (1), 1–7. doi: 10.1038/srep19075
- Tello, J., Mammerler, R., Čajić, M., and Forneck, A. (2019a). Major outbreaks in the nineteenth century shaped grape phylloxera contemporary genetic structure in Europe. *Sci. Rep.* 9 (1), 1–11. doi: 10.1038/s41598-019-54122-0
- Tello, J., Roux, C., Chouiki, H., Laucou, V., Sarah, G., Weber, A., et al. (2019b). A novel high-density grapevine (*Vitis vinifera* l.) integrated linkage map using GBS in a half-diallel population. *Theor. Appl. Genet.* 132 (8), 2237–2252. doi: 10.1007/s00122-019-03351-y
- Tenkumo, T., Aobulikasimu, A., Asou, Y., Shirato, M., Shishido, S., Kanno, T., et al. (2020). Proanthocyanidin-rich grape seed extract improves bone loss, bone healing, and implant osseointegration in ovariectomized animals. *Sci. Rep.* 10 (1), 1–14. doi: 10.1038/s41598-020-65403-4
- This, P., Jung, A., Boccacci, P., Borrego, J., Botta, R., Costantini, L., et al. (2004). Development of a standard set of microsatellite reference alleles for identification of grape cultivars. *Theor. Appl. Genet.* 109 (7), 1448–1458. doi: 10.1007/s00122-004-1760-3
- Tomić, L., Štajner, N., and Javornik, B. (2013). "Characterization of grapevines by the use of genetic markers," in *The Mediterranean genetic code-grapevine and olive*. Eds. D. Poljuha and B. Sladonja (Rijeka, Hrvatska: Intech Open Science), 1–25.
- Topfer, R., Sudharma, K., Kecke, S., Marx, G., Eibach, R., Maghradze, D., et al. (2009). The vitis international variety catalogue (VIVC): new design and more information. *Bull. IOIV* 82 (935), 45.
- Torregrosa, L., Fernandez, L., Bouquet, A., Boursiquot, J.-M., Pelsy, F., and Martínez-Zapater, J. M. (2011). Origins and consequences of somatic variation in grapevine. *Genetics Genomics Breed. Grapes* 68, 92. doi: 10.1201/b10948
- TUIK. (2019). *Turkish Statistical Institute*.
- Tyagi, A., Agarwal, R., and Agarwal, C. (2003). Grape seed extract inhibits EGF-induced and constitutively active mitogenic signaling but activates JNK in human prostate carcinoma DU145 cells: possible role in antiproliferation and apoptosis. *Oncogene* 22 (9), 1302–1316. doi: 10.1038/sj.onc.1206265
- Vondras, A. M., Minio, A., Blanco-Ulate, B., Figueroa-Balderas, R., Penn, M. A., Zhou, Y., et al. (2019). The genomic diversification of grapevine clones. *BMC Genomics* 20 (1), 1–19. doi: 10.1186/s12864-019-6211-2
- Vouillamoz, J. F., and Grando, M. S. (2006). Genealogy of wine grape cultivars: 'Pinot' related to 'Syrah'. *Hereditas* 97 (2), 102–110. doi: 10.1038/sj.hdy.6800842
- Vouillamoz, J. F., McGovern, P. E., Ergul, A., Söylemezoğlu, G., Tevzadze, G., Meredith, C. P., et al. (2006). Genetic characterization and relationships of traditional grape cultivars from transcaucasia and Anatolia. *Plant Genet. Resour.* 4 (2), 144–158. doi: 10.1079/PGR2006114
- Walker, A. R., Lee, E., and Robinson, S. P. (2006). Two new grape cultivars, bud sports of Cabernet sauvignon bearing pale-coloured berries, are the result of deletion of two regulatory genes of the berry colour locus. *Plant Mol. Biol.* 62 (4), 623–635. doi: 10.1007/s11103-006-9043-9
- Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., et al. (2009). Gplots: various r programming tools for plotting data. *R Package version* 2 (4), 1.
- Weir, B. S., and Cockerham, C. C. (1984). Estimating f-statistics for the analysis of population structure. *Evolution* 38 (6), 1358–1370. doi: 0.2307/2408641
- Wickham, H. (2011). ggplot2. *Wiley Interdiscip. Reviews: Comput. Stat* 3 (2), 180–185. doi: 10.1002/wics.147
- Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). Rmvp: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics Proteomics Bioinf* 19 (4), 619–628. doi: 10.1016/j.gpb.2020.10.007
- Yılmaz, F., Shidfar, M., Hazrati, N., Kazan, K., Özmen, C. Y., Uysal, T., et al. (2020). Genetic analysis of central Anatolian grapevine (*Vitis vinifera* l.) germplasm by simple sequence repeats. *Tree Genet. Genomes* 16 (4), 1–11. doi: 10.1007/s11295-020-01429-z
- Zdunić, G., Lukšić, K., Nagy, Z. A., Mucalo, A., Hančević, K., Radić, T., et al. (2020). Genetic structure and relationships among wild and cultivated grapevines from central Europe and part of the Western Balkan peninsula. *Genes* 11 (9), 962. doi: 10.3390/genes11090962
- Zohary, D., Hopf, M., and Weiss, E. (2012). Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. (Oxford University Press).
- Žulj Mihaljević, M., Maletić, E., Preiner, D., Zdunić, G., Bubola, M., Zyprian, E., et al. (2020). Genetic diversity, population structure, and parentage analysis of Croatian grapevine germplasm. *Genes* 11 (7), 737. doi: 10.3390/genes11070737