Kristine Margaryan[1, 2*], Boris Gasparyan[3,4], Artur Prtrosyan[3, 4], Frunz Harutyunyan[1], Reinhard Töpfer[5], Erika Maul[5]

Grapevine genetic resources of Armenia: molecular fingerprinting and phylogenetic relationship among wild and cultivated grapevine

Affiliations

[1] Research Group of Plant Genomics, Institute of Molecular Biology, National Academy of Sciences RA, Yerevan, Armenia
[2] Department of Genetics and Cytology, Yerevan State University, Yerevan, Armenia
[3] Institute of Archaeology and Ethnography, National Academy of Sciences RA, Yerevan, Armenia
[4] "Areni-1 Cave", Scientific-Research Foundation, Vayots Dzor, Areni, Armenia
[5] Julius Kühn-Institute (JKI), Institute for Grapevine Breeding Geilweilerhof, Siebeldingen, Germany

Correspondence

Kristine Margaryan*: kristinamargaryan@ysu.am, Boris Gasparyan: borisg@virtualarmenia.am, Artur Petrosyan: artur.petrosian@yahoo.com, Frunz Harutyunyan: frunz8@yahoo.com, Reinhard Töpfer: reinhard.toepfer@julius-kuehn.de, Erika Maul: erika.maul@julius-kuehn.de

Summary

Armenia is characterized by a high diversity of cultivated (*Vitis vinifera* L. subsp. *Vinifera*) and wild (*Vitis vinifera* L. subsp. *sylvestris*) grapes. The country has played a leading role in the centuries-lasting history of grapevine cultivation in the Near East. Varying climatic conditions and the existence of wild grapes lead to the formation and promotion of viticulture and winemaking, as evidenced by nearly 450 autochthonous varieties. Hundreds of unique indigenous cultivars are still preserved in old vineyards and abandoned gardens, though most of them are threatened by extinction. Wild grapes, thriving along riverbanks, climbing the rocks and embracing the trees can be found in Vayots Dzor, Tavush, Syunik provinces and in Artsakh.

With the main goal to estimate the phylogenetic relationships among Armenian wild grapes and indigenous cultivars, and evaluating the possible contribution of wild grapes to the genetic makeup of indigenous cultivars, we analyzed 79 unique cultivars and 111 putative wild plants, collected from different viticulture regions, with 24 nSSR markers.

The genetic diversity analysis conducted for wild grapes and indigenous cultivars unfolded the allelic richness of wild and cultivated gene pools and surprisingly for us revealed the absence of significant differences for all genetic parameters between the two subspecies. Moreover, the results registered for the number of different alleles (Na), effective number of alleles (Ne), and Shannon's information index (I) have shown comparatively high values for wild grapes, while the observed negative value of Fixation index (F) for indigenous cultivars mirrored an abundance of heterozygote genotypes presuming random mating. The neighbour-joining (NJ) cluster analysis indicated a clear separation between the two subspecies *vinifera* and *sylvestris* and formed two main clusters. Applied non-hierarchical horizontal clustering using Structure software assigned the 190 genotypes into two clusters. The delta K criterion (ΔK) suggested K = 2 as the optimal uppermost hierarchical level of structure. Obtained results were comparable with the NJ cluster analysis and confirmed the divergence of *sylvestris* from *vinifera*, indicating a clear separation between the two subspecies. Meanwhile, results highlighted the role of gene flow between wild grapes and cultivars through observed overlaps and admixed ancestry values. Grapevine genetic resources of Armenia can contribute to overcoming biotic and abiotic stresses and better adaptation to climate change.

Key words

wild grape, indigenous cultivar, genetic diversity, phylogeny, Armenia

Introduction

According to archaeological and historical studies grapevine domestication started during the agricultural revolution of the Neolithic period (ca. 10.000 – 5200 Cal BC) and was established in the Chalcolithic period (5200 – 3400 Cal BC), when human populations began to manage, collect and propagate *Vitis* forms to improve vine and wine production (McGovern, 2003; Zohary and Hopf, 2000; Hovhannisyan *et al.*, 2017). Domestication took place in the region spread between the Caucasus and Mesopotamian Plains, domesticating wild populations of *Vitis vinifera* subsp. *sylvestris*, considered to be the progenitor of the cultivated grapevine, the subsp. *vinifera* (Myles *et al.*, 2011; McGovern *et al.*, 2017). From the primary domestication origins, cultivated grapevines disseminated westward into neighbouring regions (North Africa and Lower Mesopotamia), reaching the Mediterranean Basin, parallel with the development of human culture. Recently, researchers applied a neural network-based machine learning method, with the idea to re-analyze the genome-wide single nucleotide polymorphism data of almost eight hundred grapevine cultivars collected from Middle Asia to the Iberian Peninsula

Submitted/accepted for publication: April 11, 2023/May 15, 2023

and from overseas regions (Nikoghosyan *et al.*, 2020; Margaryan *et al.*, 2021). Based on the results generated by self-organizing maps (SOM) portrayal genomic landscape and the different sample similarity plots were concordant with the historical knowledge and reflect the geographical distribution of grape cultivars, indicating the main pathways of grape dissemination and genome-phenotype associations about grape usage. According to data from SOM analysis, cultivated grapes occurred initially in the Caucasus, the Armenian Highlands and the so-called Fertile Crescent and then disseminated towards the Mediterranean world to the West and into the East towards Iran and the Middle East, Afghanistan and India. The northern and southern ways into the west agree with the distribution of settlements of Greeks and Phoenicians, respectively.

The large-scale research launched recently elucidate grapevine evolution and domestication history with 3,525 cultivated and wild European grapevines. According to the study domestication occurred concurrently about 11,000 years ago in Western Asia and the Caucasus parallel to yield table and wine grapes respectively (Yang Dong *et al.*, 2023).

During dissemination, the grapevine increased its genetic diversity and variability due to the contribution of multiple genetic pools and continuing human selection. In the enrichment of grape genetic heterogeneity, the important role belongs to wild populations, introgressed hybrids between varieties and local wild forms, *de novo* domesticated forms from native wild grapevines and somatic mutants of cultivars as the main source of phenotypic variation generated diversity without modifying the identity of the variety (Di Gaspero and Testolin, 2013).

The wide range of altitudinal variation and different climate zones in Armenia are the main drivers of extremely rich plant biodiversity and includes many regionally endemic, relict and rare species. The country is an important center of endemism for wild relatives of economically important crops, including the grapevine as the most emblematic plant. Due to its strategic position, Armenia has had a key role in the spreading of grapevine, viticulture and winemaking practices through the centuries. The country is a unique grapevine diversity hotspot where the viticulture and wine industry are dating back to the end of the V to the beginning of the IV Millennium BC (4100-3800 Cal BC). The oldest wine making complex discovered in Armenia is located near the Areni village of Vayots Dzor Province, in the Areni-1 cave (Hobosyan *et al.*, 2021). The excavated industrial complex is the best-preserved monument, which is a testament to the six thousand years of wine-making tradition in the area. The process of grapevine domestication was linked with important modifications in architecture and biology of the grapevine in comparison with other crops (This *et al.*, 2006). The phenotypic gap between subspecies associated with domestication is the development of hermaphroditic flowers in *vinifera*, the increased number of berries per cluster, the enlargement of berry size and seedlessness in table grapes. The biology of grapes changed dramatically to guarantee greater yield, higher sugar content for better fermentation and more regular production. Debates about whether these changes occurred through sexual reproduction and natural or targeted human selection, or via mutation, selection and following vegetative propagation still remain unclear (This *et al.*, 2006).

The understanding of the importance of the protection and conservation of genetic resources of the grapevine wild ancestor *Vitis vinifera* L. subsp. *sylvestris* is very high for several reasons. Nowadays, the wild grape population has become relict due to several forms of human disturbance such as habitat destruction and fragmentation, irregular management of the natural environment, pathogen spread, which has increased in the last decades, and a demanding reproductive strategy (Margaryan *et al.*, 2019*).*

Studies on wild grapes reinforced in parallel with advanced molecular technologies, the ultimate goals of preserving its biodiversity, clarifying its taxonomic status and identifying traits of interest for the breeding program. Wild grapes, thriving along riverbanks, climbing the rocks and embracing the trees still can be found in Armenia.

Viticulture in Armenia was most extensive in the 20th century and about 850 varieties were preserved in the first grapevine national collection (Fig. 1). The collection was established in

Fig. 1: The first Armenian grapevine collection, after Beketovskiy D.A., 1968.

the 1950s at the Institute of Wine-Making and Fruit-Growing (Margaryan *et al.*, 2021). Among conserved varieties, 400-450 were autochthonous grapevines, mainly the result of spontaneous hybridization. Due to different socio-economic reasons and the collapse of the Soviet Union, the collection was destroyed in the early 1990s and only in 2016 within the nationwide program supported by Food and Agriculture Organization (FAO) the new collection was established conserving almost 300 grapevine accessions.

Today, about half of autochthonous varieties registered earlier in different bibliographic resources are irreversibly lost due to various reasons or existing as single vines and are therefore seriously threatened with extinction. Unfortunately, the globalization of the wine market and the factor of variety-oriented wine labelling have led to the decline of the genetic diversity of wine grape varieties. The ignorance of minor autochthonous varieties having only local importance resulted in an alarming reduction since only thirty to thirty-five of 400-450 native grape varieties are used in wine and brandy production. In this context, it is essential to develop and use effective strategies for the systematic analysis of the grapevine diversity, and efficient use and maintenance of grape germplasm resources (Margaryan *et al.*, 2021).

The proper number of Armenian varieties is not known and there is often uncertainty about their precise number since synonyms and homonyms occur. Assessment of genetic diversity and verification of the genetic relatedness among *sylvestris* and *vinifera* accessions, combined with parentage analysis, are the first steps in a dissection of grape genetics. For confirmation of variety identity, the symbiotic approach combining molecular and morphological characterization for each genotype was carried out. Genetic characterization was performed by 24 nuclear simple sequence repeats (nSSR), which are the most utilized due to well-documented advantages. Phenotypic studies were done using the "Multi Crop Passport Descriptors" (MCPD) of Bioversity. MCPD-data provide basic information about the accession, including accession name, accession number, which is a unique code assigned by the curator of the collection, berry colour, provenance, donor, *etc*. These studies involved also the comparison of morphological features for cultivated varieties with existing bibliographies and online databases, such as the *Vitis* International Variety Catalogue (*V*IVC) (https://www.vivc.de/) and Réseau Français des Conservatoires de Vigne (https://bioweb.supagro.inra.fr/collections_vigne/SearchS.php).

The presented research was part of the most representative and comprehensive analysis of Armenian grape germplasm started in 2017 in close cooperation with the Institute for Grapevine Breeding, JKI intending preservation, promotion and prominence of native grape germplasm towards recovering the untapped biodiversity and breeding potential of *V. vinifera* in Armenia (Margaryan et *al.*, 2021).

The main idea of the proposed research was to estimate the phylogenetic relationships between Armenian wild grapes and indigenous cultivars, to evaluate the genetic diversity of two subspecies and to estimate the possible contribution of wild grapes to the genetic makeup of indigenous cultivars.

Material and Methods

Plant material

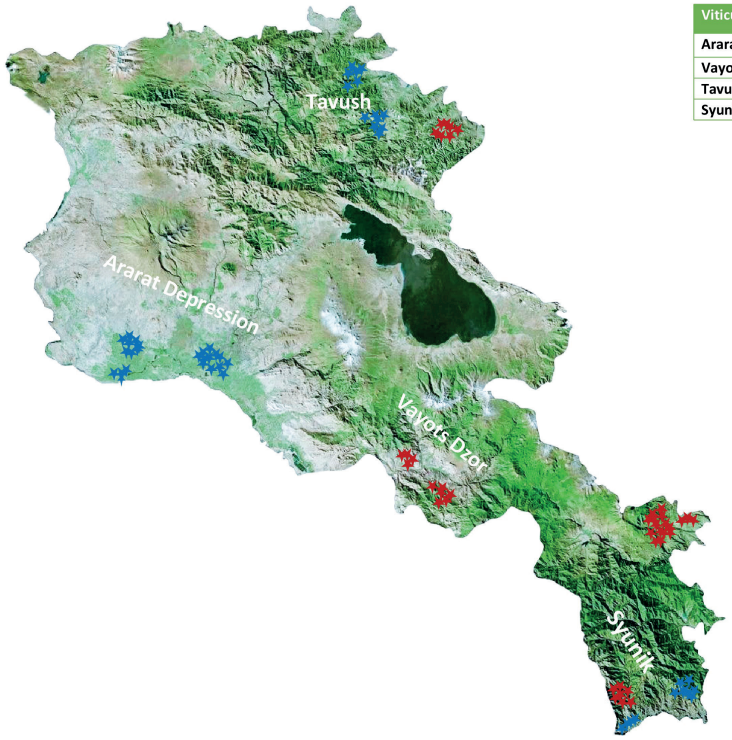
Significant efforts were carried out to recover and identify local minor grapevine germplasm and wild grapes in traditional viticulture regions of Armenia: Ararat, Aragatsotn, Vayots Dzor, Tavush, Syunik, and Artsakh during the vegetation and harvest period. The nationwide survey mainly focused on vineyards established at the beginning of the 20th century and earlier; some of them were totally out of cultivation for a long time. Family gardens were included, as well as a few small private collections located in Ararat Depression and Tavush. The material was collected through the support of local farmers and industry members. Accession designations and MCPD-data were recorded. Sometimes grapevines did not have a varietal name and the generic names were used for the accessions, referring, for example, to a morphological trait, like grape colour or shape, farmer's name and planted area, village. GPS coordinates and elevations of the sampled accessions were registered. *V. sylvestris* plants were collected in their natural habitat: riverbanks, climbing the rocks and growing on trees. Each putative wild candidate was analyzed morphologically and only those that met the basic phenotypic characteristics of wild grapevines were subjected to further genetic analysis. *V. sylvestris* plants were sampled at their natural habitats on four different locations (Fig. 2).

According to our observations the tip of the young shoot of wild plants was fully open without anthocyanin pigmentation for all accessions collected from different provinces, while the shape of mature leaves was diverse from three to five lobes, sometimes non-lobed. The size of the mature leave usually was medium and small and the petiole sinus is half-opened. For wild plants collected from Lori region the presence of prostrate hairs between main veins was detected.

A total of seventy-nine endangered autochthonous grapevine varieties and one hundred eleven wild grapes were collected for molecular fingerprinting.

DNA extraction and microsatellite analysis

Total genomic DNA was extracted from 100 mg of young leaf tissue after grinding with MM 300 Mixer Mill system (Retsch, Haan, Germany) and stored at −80 °C until use. DNA extraction was performed employing the DNeasy 96 plant mini kit (QIAGEN, Dusseldorf, Germany) following the manufacturer's protocol. DNA concentration and quality were checked by spectrophotometric analysis and electrophoresis in a 1% agarose gel. Microsatellite fingerprinting of genotypes was performed on 24 microsatellite loci (nSSRs) well distributed across the 19 grape chromosomes as previously described (Laucou *et al.*, 2011; Sefc *et al.*, 1999) (*i.e.*, VVS2, VVMD5, VVMD7, VVMD21, VVMD24, VVMD25, VVMD28, VVMD27, VVMD32, four of the VrZAG series (VrZAG62, VrZAG79, VrZAG67, VrZAG83), VMC4f3.1, VMC1b11 and nine of the VVI series VVIb01, VVIn16, VVIh54, VVIn73, VVIp31, VVIp60, VVIv37, VVIv67, and VVIq52). Nine polymorphic microsatellite markers proposed by the GrapeGen06 project: VVMD5, VVMD7, VVMD25, VVMD27, VVMD28, VVMD32, VVS2,



Viticultural regions	*V. v.* L. subsp. *vinifera*	*V. v.* L. subsp. *sylvestris*
Ararat Depression	29 cultivars	
Vayots Dzor	25 cultivars	10 genotypes
Tavush	15 cultivars	12 genotypes
Syunik, Artsakh	10 cultivars	89 genotypes

Fig. 2: Geographic distribution of the sampled V. v. L. subsp. vinifera and V. v. L. subsp. sylvestris populations in Armenia.

VrZAG62, and VrZAG79 were used for comparison of genetic profiles with the SSR-marker database of the Julius Kühn-Institut (JKI), maintaining about 8,000 genetic profiles from distinct sources. Fingerprints from the European *Vitis* Database produced during the European project GrapeGen06 and data from COST Action FA1003 were used to find corresponding profiles (Maul *et al.*, 2012; Bacilieri and This, 2010; Maul *et al.*, 2015). Fifteen additional markers were applied for parent-offspring analysis.

For fragment length determination by capillary electrophoresis on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Life Technologies, Waltham, MA, USA), all forward primers were 5'-labelled with a fluorescent dye (FAM, HEX, TAMRA, ROX and PET). The combination of markers with different labels and diverse fragment lengths allows one to perform the polymerase chain reaction (PCR) and grouped markers in seven multiplex pools, comprising two to five SSR markers characterized by similar annealing temperatures. The 2x KAPA2G Fast PCR Kit (Merck, Düren, Germany) was used to set up 5 µL reaction mixtures containing master mix, 100 pmol of each primer and 1 ng of template DNA. GeneAmp PCR system 9700 thermal cycler (Applied Biosystems, Schwerte, Germany) was used for the amplification starting with 3 min initial denaturation at 95 °C, followed by 30 cycles for 30 s. A final extension was performed at 72 °C for 7 min. 1 µL of the PCR product was used for fragment length determination and the results were processed with GeneMapper 4.0 software (Applied Biosystems, Life Technologies, Waltham, MA, USA) recorded in base pairs. Allele size was determined by comparing fragment peaks with the internal size standard, using the Microsatellite default method for size calling with SSR and the expected repeat size. To correct the amplification shifts among different multiplexes, SSR profiles were adapted by including in each PCR amplification run the DNA of standard cultivars 'Cabernet franc' and 'Muscat à petits grains blancs'.

Data analysis

The genetic diversity among wild and cultivated grapevine groups was estimated. The standardized nSSR genotyping data were used to determine the number of different alleles (Na), the number of effective alleles (Ne), Shannon's Information Index (I), observed heterozygosity (Ho), expected heterozygosity (He, fixation index (F) referred to as the inbreeding coefficient and private alleles (PA). The allele frequency for each nSSR locus was calculated as well. GenAlEx software version 6.5 was used to compute genetic diversity statistics for each nSSR locus (Nei, 1973; Peakall and Smouse, 2006). Clustering was performed by MEGA 7 software, version 7.0.26, which was used to generate a distance tree by the Neighbor-joining (N-J) hierarchical clustering method using the codominant genotypic distances between all pairwise combinations calculated by the GenAlEx 6.5 (Kumar *et al.*, 2016; Saitou and Nei, 1987). Principal coordinate analysis (PCoA) was used to display genetic divergence among samples using codominant genotypic distances computed in GenAlEx 6.5 (Kalinowski *et al.*, 2007).

The admixture model in Structure 2.3.4 (Pritchard *et al.*, 2000) was employed to infer the number of genetic populations (K) existing in the samples and to assign genotypes to populations of origin, with no prior information. The Structure configuration was set to ignore population information and use an admixture model with correlated allele frequencies. Various numbers of putative populations (K) were test-

ed, ranging from 1 to 10. Burning time and replication number were set to 100,000 and 100,000, respectively, in each independent run with 10 iterations. The choice of the most likely number of clusters (best K) was evaluated following the ad hoc statistic delta K (ΔK) as described by Evanno *et al.*, 2005) using Structure Harvester (Earl and Vonholdt, 2012). The Structure bar plot was visualized by running the clump file (K = 2) obtained by Structure Harvester, in Structure Plot v 2.0 (Ramasamy *et al.*, 2014). Structure ancestry Q values for each analysed individual were calculated with the highlighted values of Q > 0.75 representing reliable ancestry assignment to its own cluster.

Results and Discussion

Genetic data from 24 nSSR loci and across 190 grapevine accessions, originating from different viticulture regions of Armenia and representing both subspecies of *V. vinifera* (subsp. *vinifera and subsp. sylvestris*), were used in this study. Microsatellite profiles were used to calculate statistical indices and compute the genetic diversity of the 111 wild and 79 cultivated genotypes (Tables 1 and 2). The range of allele size (Ra), number of different alleles (Na), effective number of alleles (Ne), Shannon's information index (I), observation heterozygosity (Ho), expected heterozygosity (He), and fixation index (F) were calculated to assess the genetic diversity of wild and cultivated grapevines of Armenia. The high number of different alleles determined for *V. sylvestris* and *V. vinifera* (respectfully 266 and 265) proves the high degree of genetic variability. The mean number of alleles per locus (Na) was 11.083 for wild and 11.042 for cultivated accessions. The number of effective alleles (Ne) for wild samples ranged from 2.031 (VVIn16) to 8.767 (VVIv37), with an overall mean of 5.259. The mean Shannon's Information Index (I) value for the wild samples was slightly higher than that for the cultivars (1.823 vs. 1.779).

The observed and expected heterozygosity (Ho, He) are considered to evaluate the genetic variability among analyzed genotypes. Observed and expected heterozygosity for *V. sylvestris* was least for VVIn73 and VVIn16; the largest Ho was for VVIp31 while the largest He was for VVIv37. VVIp31 is considered the most informative marker in the wild population. The expected heterozygosity (He) values varied between 0.508 (VVIn16) and 0.886 (VVIv37), with an average of 0.780.

Table 1: Descriptive statistics and genetic diversity of the 111 wild grapes at 24 microsatellite loci

Locus	Ra (bp)	Na	Ne	I	Ho	He	F
VVS2	125-157	12	6.153	2.006	0.743	0.837	0.113
VVMD5	228-248	10	6.494	2.030	0.716	0.846	0.154
VVMD7	235-263	12	7.439	2.168	0.773	0.866	0.107
VVMD25	237-271	10	3.051	1.441	0.609	0.672	0.094
VVMD27	176-198	11	5.766	1.956	0.748	0.827	0.095
VVMD28	228-282	16	6.783	2.237	0.781	0.853	0.084
VVMD32	240-292	14	6.954	2.200	0.791	0.856	0.076
VrZAG62	188-204	7	5.107	1.737	0.829	0.804	-0.031
VrZAG79	237-261	12	6.221	2.068	0.802	0.839	0.045
VVIv67	348-401	15	5.082	2.003	0.769	0.803	0.043
VrZAG67	122-159	12	5.164	1.933	0.775	0.806	0.039
VrZAG83	188-201	4	3.821	1.363	0.700	0.738	0.052
VVIn16	147-155	4	2.031	0.878	0.514	0.508	-0.012
VVIn73	258-272	5	2.091	0.887	0.482	0.522	0.076
VVIp60	302-331	13	5.333	1.961	0.709	0.812	0.127
VVMD24	206-218	6	4.122	1.562	0.764	0.757	-0.008
VVMD21	244-267	9	4.772	1.689	0.624	0.790	0.211
VMC4f3.1	163-217	17	7.108	2.340	0.807	0.859	0.060
VVIb01	289-319	9	3.082	1.393	0.655	0.675	0.031
VVIh54	139-179	15	6.455	2.085	0.706	0.845	0.164
VVIq52	70-86	9	3.030	1.465	0.636	0.670	0.050
VVIv37	144-180	13	8.767	2.289	0.731	0.886	0.174
VMC1b11	167-203	16	5.028	1.976	0.766	0.801	0.043
VVIp31	157-195	15	6.352	2.077	0.835	0.843	0.009
Total		*266*					
Min.		*4*	*2.031*	*0.878*	*0.482*	*0.508*	*-0.031*
Max		*17*	*8.767*	*2.340*	*0.835*	*0.886*	*0.211*
Mean		*11.083*	*5.259*	*1.823*	*0.719*	*0.780*	*0.075*

Table 2: Descriptive statistics and genetic diversity of the 79 non-redundant cultivars at 24 microsatellite loci

Locus	Ra (bp)	Na	Ne	I	Ho	He	F
VVS2	125-155	14	7.966	2.211	0.899	0.874	-0.028
VVMD5	226-248	10	6.201	1.990	0.886	0.839	-0.056
VVMD7	233-259	12	5.296	1.891	0.844	0.811	-0.041
VVMD25	239-267	8	4.760	1.696	0.886	0.790	-0.122
VVMD27	176-198	8	4.313	1.616	0.835	0.768	-0.088
VVMD28	218-282	17	6.060	2.178	0.859	0.835	-0.029
VVMD32	240-292	13	4.586	1.931	0.785	0.782	-0.004
VrZAG62	186-206	11	6.424	2.008	0.861	0.844	-0.019
VrZAG79	237-261	12	5.088	1.917	0.835	0.803	-0.040
VVIv67	348-401	14	3.279	1.752	0.734	0.695	-0.056
VrZAG67	122-159	16	7.606	2.256	0.873	0.869	-0.006
VrZAG83	188-201	5	2.653	1.097	0.615	0.623	0.012
VVIn16	144-157	6	2.508	1.173	0.684	0.601	-0.137
VVIn73	258-272	7	2.142	1.069	0.538	0.533	-0.010
VVIp60	306-332	13	3.492	1.640	0.756	0.714	-0.060
VVMD24	204-220	9	3.839	1.612	0.747	0.740	-0.010
VVMD21	244-267	7	3.364	1.401	0.835	0.703	-0.189
VMC4f3.1	163-217	18	7.885	2.331	0.848	0.873	0.029
VVIb01	289-313	9	3.275	1.441	0.785	0.695	-0.130
VVIh54	139-177	13	4.943	1.964	0.848	0.798	-0.063
VVIq52	70-82	6	3.357	1.332	0.772	0.702	-0.100
VVIv37	150-178	13	8.310	2.221	0.896	0.880	-0.019
VMC1b11	167-205	13	4.566	1.843	0.821	0.781	-0.051
VVIp31	171-193	11	7.270	2.128	0.937	0.862	-0.086
Total		*265*					
Min.		*5*	*2.142*	*1.069*	*0.538*	*0.533*	*-0.189*
Max		*18*	*8.310*	*2.331*	0.937	*0.880*	*0.029*
Mean		*11.042*	*4.966*	*1.779*	*0.808*	*0.767*	*-0.054*

Overall mean observed heterozygosity was lower than the expected heterozygosity. The observed heterozygosity (Ho) values ranged from 0.482 (VVIn73) to 0.835 (VVIp31), with an overall mean of 0.719. The estimated Fixation index, a measure of reduction in heterozygotes (also called inbreeding coefficient F_{IS}) in the wild genotypes showed a positive value of 0.075 ± 0.012, pointing to a heterozygote deficiency in the wild group.

The article includes an analysis of 79 non-redundant grapevine varieties for which the assessment of trueness to type was done by an approach combining molecular fingerprinting, morphological description, and in-depth bibliographic studies. For each identified variety the genetic profile was compared with almost 8,000 fingerprints documented in the JKI-SSR-marker database (European *Vitis* Database) and genetic profiles generated during COST Action FA1003 and bibliography. Armenian (published by Tumanyan, 1947; Poghosyan, 1962; Poghosyan, 1981; Melyan, 2019) and Russian ampelographies (published by Frolov-Bagreev *et al.*, 1946–1956; Kartavchenko and Blagonravov, 1963–1970; Golodriga *et al.*, 1984), as well as Caucasus and Northern Black Sea Region Ampelography (published by Maghradze *et al.*, 2012), were used in the study as the main important sources.

The number of alleles per nSSR locus for cultivated grapevines ranged from 5 (VrZAG83) to 18 (VMC4f3.1). The effective number of alleles, respecting alleles that occur at a relevant frequency within the sample, ranged from 2.142 for locus VVIn73 to 8.310 for locus VVIv37, with a mean of 4.966. The following VVMD28 and VrZAG67 loci displayed high Ne values as well. The highest Shannon's information index (I) was observed in VMC4f3.1 locus (2.331) and lowest in VVIn73 (1.069). Shannon's information index is an important parameter mirroring the level of polymorphism. The observed heterozygosity (Ho) for cultivars ranged from 0.538 (VVIn73) to 0.937 (VVIp31), with an overall mean of 0.808. The expected heterozygosity (He) values ranged between 0.533 (VVIn73) to 0.880 (VVIv37), with an average of 0.767. The overall mean of observed and expected heterozygosity was slightly different. The estimated Fixation index (F) as an indicator of inbreeding ranged between −0.189 (VVMD21) to 0.029 (VMC4f3.1) with a mean value of -0.054. The observed negative F values indicated an abundance of heterozygote genotypes presuming random mating.

Fifty-three private alleles (PA) were detected at 18 out of 24 nSSR markers in the studied set of *sylvestris* and fifty-two PA were detected at 17 out of 24 nSSR markers in *vinifera* (Table 3). The number of PA in both of subspecies are practically the same. The highest number of private alleles was identified for VVMD28, VVIv67 and VMC4f3.1 marker in wild populations and VMC4f3.1 have shown the greatest number of PA for cultivated grapevines as well.

The neighbour-joining (N-J) distance tree was constructed to investigate the genetic relationship among the 79 non-redundant cultivars and 111 wild grapes (Fig. 3). The hierarchical clustering of 190 unique genotypes distinguished two major clusters (CI, CII) with sub-clusters and showed clear grouping of *V. vinifera* and *V. sylvestris.* Anyway, some wild genotypes clustered with the cultivated samples and *vice versa*.

Cluster I with two main sub-clusters encompasses 71 wild accessions originating from the Syunik region and Artsakh. The sub-cluster 1 of Cluster I is a blend of wild grapes collected from both regions, while sub-cluster 2 groups wild grapes only from Syunik. In this second sub-cluster surprisingly for us, 'Sveni' was grouped with wild grapes. 'Sveni' is a registered variety and is characterized in the second book of the Armenian Ampelography published in 1981. Only a few plants were found in the old vineyards of Dzorashen, Aygedzor and Khndzoresk villages in Goris, Syunik region. According to the records of ampelographers describing varieties for the first time, the plants were almost 100 years old with female phenotype and varied-sized cylindrical-conical shaped bunches. The variety is black-berried and covered with a dense layer of wax. Further molecular fingerprinting and following parentage analysis detected two wild genotypes S 54 and S 71 as potential parents, originating from Syunik. This fact was confirmed also by Illumina sequencing and the status of 'Sveni' based on whole genome data was assigned as *sylvestris* (article submitted for publication). Cluster II includes all autochthonous varieties and some accessions collected as putative wild grapes. The pre-selection of true-to-type *sylvestris* collected during surveys was done based on the phenotypic characterization of leaves, flower sex (during flowering period, when it was possible) and berries.

According to the clustering results, the first sub-cluster of Cluster II grouped 20 genotypes including three cultivars 'Karch mat', 'Khatuni' and 'Nelsoni Tsaghkavan' which was named by us, because this formerly unknown variety was found in

Table 3: List of private alleles and allele frequencies in 190 sample set of *V. sylvestris* and *V. vinifera*

SSR locus	*V. sylvestris*	*V. vinifera*
VVS2	157(0.009)	131(0.006), 139(0.006), 147(0.006)
VVMD5	244(0.124)	226(0.025)
VVMD7	257(0.009), 261(0.009), 263(0.132)	233(0.019), 241(0.006), 251(0.013),
VVMD25	237(0.032), 269(0.005), 271(0.005)	247(0.006)
VVMD27	188(0.005), 192(0.023), 193(0.018)	-
VVMD28	228(0.038), 252(0.010), 254(0.071), 264(0.014), 276(0.076)	218(0.058), 226(0.006), 238(0.006), 242(0.013), 260(0.013), 280(0.013)
VVMD32	246(0.041)	-
VrZAG62	-	186(0.013), 192(0.013), 198(0.006), 206(0.006)
VrZAG79	-	-
VVIv67	349(0.014), 363(0.014), 374(0.032), 380(0.042), 382(0.023)	350(0.013), 356(0.006), 361(0.006), 391(0.013)
VrZAG67	130(0.032)	124(0.032), 147(0.006), 152(0.006), 157(0.006)
VrZAG83	-	180(0.006)
VVIn16	-	-
VVIn73	-	264(0.058)
VVIp60	302(0.036), 304(0.023)	314(0.019), 332(0.045)
VVMD24	-	204(0.057), 212(0.006), 220(0.006)
VVMD21	251(0.179), 255(0.005)	-
VMC4f3.1	169(0.005), 178(0.005), 207(0.064), 211(0.023), 213(0.023)	175(0.006), 176(0.013), 183(0.006), 190(0.013), 196(0.006), 208(0.006)
VVIb01	317(0.055), 319(0.055)	305(0.006), 313(0.006)
VVIh54	141(0.018), 161(0.005), 173(0.005), 179(0.009)	159(0.025), 169(0.013)
VVIq52	72(0.005), 84(0.032)	-
VVIv37	144(0.005), 148(0.074), 180(0.005)	162(0.006), 170(0.006), 172(0.013)
VMC1b11	177(0.019), 197(0.009), 199(0.093), 201(0.009), 203(0.005)	171(0.032), 205(0.006)
VVIp31	157(0.005), 169(0.005), 191(0.014), 195(0.041)	-

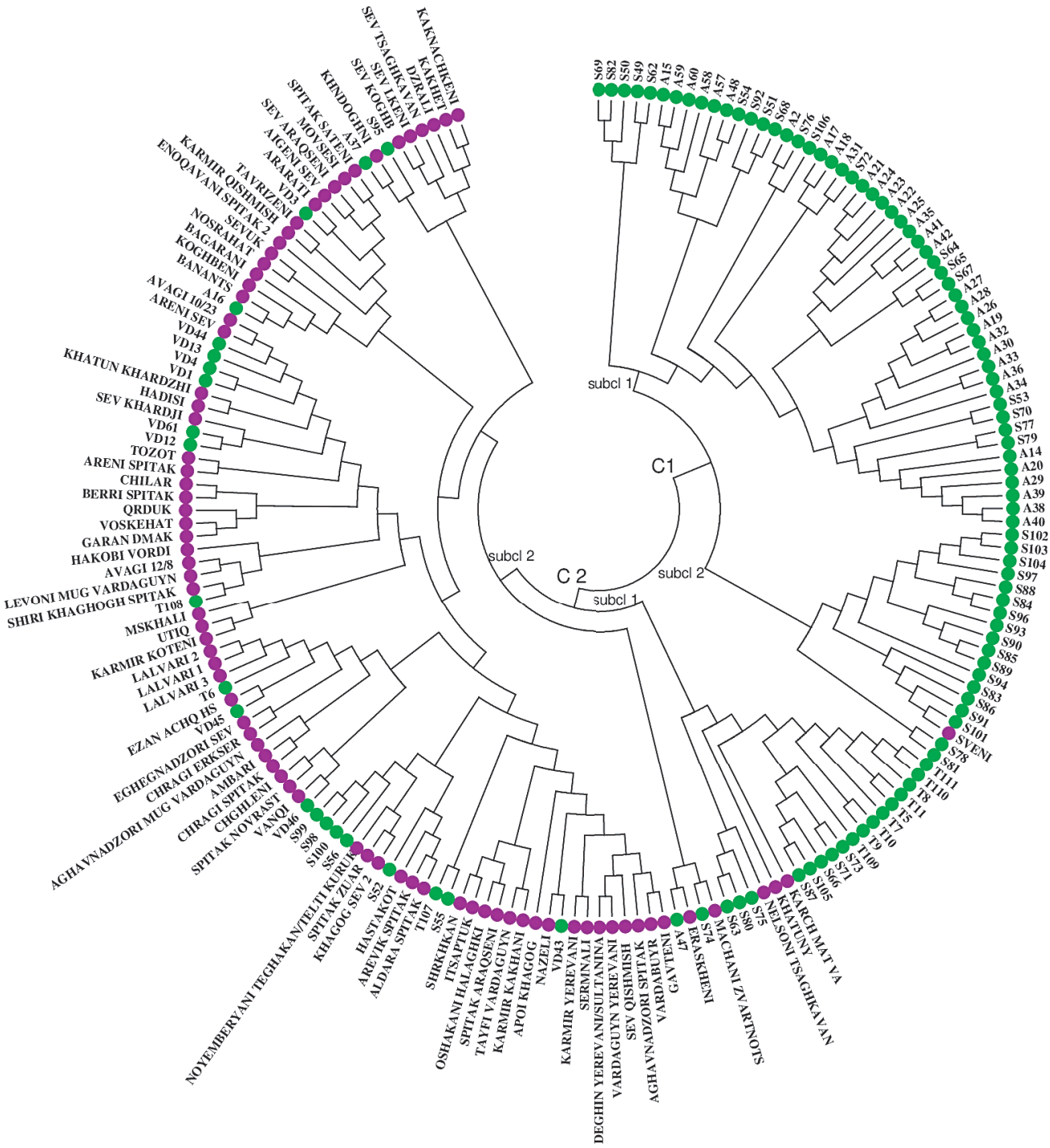


Fig. 3: Neighbor-joining dendrogram showing genetic distance among Armenian Vitis vinifera wild vines and cultivars, based on 24 nSSR loci. Armenian autochthonous varieties are marked with violet and wild grapes with green colour.

Tsaghkavan village in Tavush. In this sub-cluster, 9 accessions were collected as wild plants from the forests of the Tavush region (T111, T110, T8, T11, T5, T7, T10, T9 and T109). The cultivars 'Khatuni' and 'Karch mat VA' have shown a close genetic relationship with samples collected as wild plants from the Syunik region. In the previous publication (Margaryan *et al.*, 2021) the accession collected and described as 'Karch mat' showed a matching genetic profile to 'Tozot', which is one of the emblematic wine grapes of the Vayots Dzor region. 'Karch mat' is described as a distinct variety, however shares close morphologic pattern with 'Tozot' such as circular, five-lobed leaves, conical, dense or very dense bunches and black, slightly elliptic or ovate berries covered by a dense or middle dense layer of wax. According to the first Armenian Ampelography (published in 1947), this variety was found by E. Aslanyan in old vineyards in the villages Artashat and Dvin located in the Ararat Depression and the name of the variety is associated with plant morphology since 'Karch mat' translated means a short plant of weak vigour. Later more detailed characterization was done by Poghosyan and the variety is registered also in the second volume of the Armenian Ampelography published in 1962. During a survey in 2020 of the old

vineyard located in the village Verin Artashat, the accession named 'Karch mat' was discovered. In contrast to the genotype collected from the Vayots Dzor region and bibliographic description, this sample has a female phenotype. Further molecular fingerprinting revealed that the sample 'Krach mat VA' matches the genetic profile of 'Pinot Noir' with only differences related to flower sex. Thus, the final decision related to the correct identity of 'Karch mat' needs further clarification. A majority of the identified varieties were grouped in Cluster II including also native, minor and unknown genotypes. In Cluster II the ancient indigenous grape varieties 'Dzrali', 'Sev Lkeni', 'Sev Koghb' originating from Tavush were grouped with unknown local varieties 'Kaknachkeni' and 'Sev Tsakhgavan' named by us.

Cluster II encompasses native varieties of the Vayots Dzor region such as 'Areni Sev', 'Areni Spitak', 'Khardji Sev', 'Tozot', grouping also some unknown varieties coded as Avagi 10/23, Avagi 12/8. These varieties were collected from vineyards more than 120 years old, located in the village Artabuynq. It is important to note, that the samples VD 44, VD43, VD 4, VD 1, VD 61 and VD 12 were collected as putative wild grapes from Gnishikadzor and mountainous areas of the Noravanq monastery in the Vayots Dzor region. They grouped with cultivars from the same geographic region. The NJ clustering was carried out without considering the geographic origin of the genotypes. The cultivated grapevines demonstrated a clear structured positioning of accessions by ancestry, diversity and putative geographic origin.

In order to identify the structure of populations and the correlations among subspecies and genotypes, a non-hierarchical PCoA based on the genetic distance matrix was used to analyze the relationships between wild and cultivated grapevines as revealed by nSSR markers (Fig. 4). The first two principal axes explain only 17.61% of the total variation (12.61 and 5%, respectively). Genetically similar or related genotypes were highly correlated and clustered together, forming two main groups. The applied PCoA revealed overlap between *vinifera* and *sylvestris*, pointing out possible gene flow from *vinifera* to *sylvestris* and *vice versa*.

The third method applied to evaluate the relationship among genotypes was non-hierarchical clustering algorithm implemented in the program STRUCTURE. Horizontal clustering assigned the 190 genotypes (111 *V. sylvestris* and 79 *V. vinifera*) into two clusters. The optimum value for the *ad hoc* number, based on the second order rate of probability of the likelihood function respecting to Delta K, was attained for K = 2. Obtained data were roughly comparable with the N-J cluster analysis and PCoA results (Fig. 5). Both groups were clearly separated, but also showed some admixture, corresponding to the overlap observed in the N-J and PCoA analysis. The accessions with Q < 75 were considered accessions with admixed origin.

Based on the results of the Q matrix value, which is an estimated membership coefficient for each genotype in each K cluster, revealed that 21 wild accessions out of 111 were far less likely to be assigned to the *sylvestris* cluster. The majority of these genotypes with mixed ancestry were collected from the Vayots Dzor region. It is important to note that phenotypically, these genotypes shared traits similarity with *sylvestris* such as small and very loose bunches, small rounded berries with violet-black skin color, and small leaves mainly with three lobes. Admixed genotypes are almost 8% (15 samples) of the studied set and among them are three cultivars: 'Karch mat VA', 'Enoqavani spitak 2' and 'Nelsoni Tsaghkavan'. For the last two samples the identity is still unclear and both of them are included in the set of "unknown" grapes. Twelve out of fifteen admixed individuals originated from the Syunik region and are most likely feral hybrids of *sylvestris* and *vinifera*, which can have a great contribution in breeding programs.

The applied approach using NJ, PCoA, and STRUCTURE analyses convinced a clear separation of the two subspecies. Generally, both *vinifera* and *sylvestris* showed high assignment to their cluster, confirming the representativeness of the studied set. The genotypes located in the transition zone in NJ, the overlapping zone in PCoA or the admixed part in STRUC-

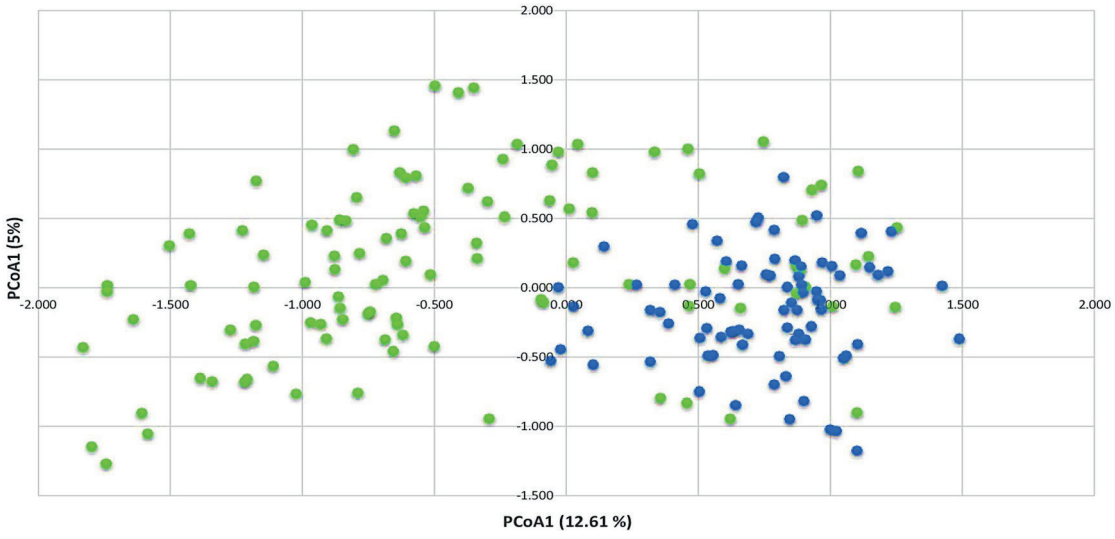


Fig. 4: Principal Coordinates Analysis (PCoA) of the 190 wild (in blue colour) and cultivated varieties (in green colour) represented by two axes. Analysis based on 24 nSSR loci via distance matrix with data standardization.

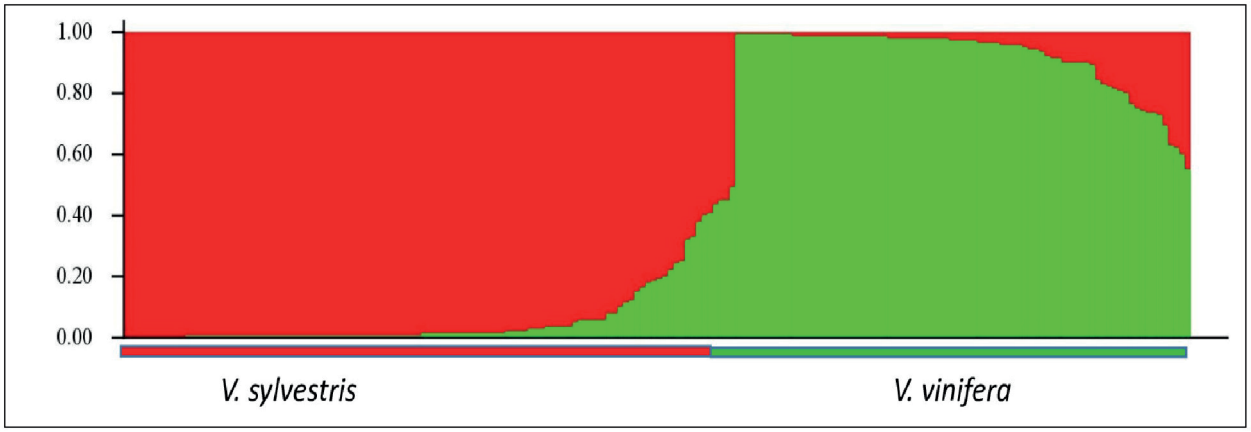


Fig. 5: Graphic presentation of the population structure of 190 grapevine accessions. Each accession is represented by a single vertical bar divided into K color segments representing its proportions in the two inferred genetic clusters using STRUCTURE software. Wild accessions grouped into a population represented by green, while cultivated accessions grouped into a population represented by red.

TURE suggest close genetic relationships due to crosses and common gene pool of the two subspecies. Based on obtained results, wild and cultivated grapevines involved in the study demonstrated a high level of polymorphism and heterozygosity across 24 nSSR loci and significant genetic diversity was detected within and between two subspecies. In accordance to the previous studies (Margaryan *et al.*, 2019; Riaz *et al.*, 2018), higher level of heterozygosity in *sylvestris* was expected, which is explained also by its obligate out-crossing nature.

Conclusions

The study of genetic relationships among the two subspecies of *Vitis vinifera* evidenced genetic relatedness between wild and cultivated grapes in Armenia. The applied hierarchical and non-hierarchical clustering methods differentiated between *sylvestris* and *vinifera*, but also demonstrated existence of gene flow between the wild and cultivated grapevines through overlaps and presence of admixed ancestry values. High levels of genetic diversity demonstrated by the effective number of alleles and richness of private and new alleles, mirrored the existence of significant diversity both within and between the subspecies suggesting that Armenia is an important center of grape biodiversity.

Conflicts of interest

The authors declare that they do not have any conflicts of interest.

Acknowledgements

This study was supported by the Science Committee, Ministry of Education, Science, Culture and Sports RA in the frame of the projects 20APP-4E007, 21T-1F076 and by ANSO-CR-PP-2020-04-A project.

The experiments and analyses were conducted at the Institute of Grapevine Breeding, JKI. We would like to thank laboratory technicians for assistance. We would like to thank the Areni-1 Cave Consortium ("Areni-1 Cave" Scientific-Research Foundation, Armenia and the "Gfoeller Renaissance Foundation USA") and Vine and Wine Foundation of Armenia for contributing to field trips. Our special acknowledgements are dedicated to our colleague Gagik Melyan, who, unfortunately, passed away.

References

Bacilieri, R., This, P., 2010: GrapeGen06, an European Project for the Management and Conservation of Grapevine Genetic Resources. Available online: http://www1.montpellier.inra.fr/grapegen06/.

Di Gaspero, G., Testolin, R., 2013: Grapevine Genomics and Phenotypic Diversity of Bud Sports, Varieties and Wild Relatives. In: Poltronieri, P., Burbulis, N., Fogher, C. (Eds.): *From plant genomics to plant biotechnology*. Woodhead Publishing Limited, DOI: 10.1533/9781908818478.149.

Dong, Y., Duan, Sh., Xia,Q., Liang,Zh., Dong, X., Margaryan, K., 2023: Dual domestications and origin of traits in grapevine evolution. Science 379 (6635), 892-901. DOI: 10.1126/science.add8655.

Earl, D.A., Vonholdt, B.M., 2012: STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4, 359-361. DOI: 10.1007/s12686-011-9548-7.

Evanno, G., Regnaut, S., Goudet, J., 2005: Detecting the number of clusters of individuals using the software structure: A simulation study. Molecular Ecology 14, 2611-2620. DOI: 10.1111/j.1365-294X.2005.02553.x.

Frolov-Bagreev, A.M., Negrul, A.M., Blagonravov, P.P., 1946-1956: Ampelogafiya SSSR I–VI (Ampelography of SSSR I–VI), Pishchepromizdat: Moskva (Moscow), Russia. (In Russian).

Golodriga, P.Y., Panarina, A.M., Troshin, L.P., 1984: Ampelogafiya SSSR (Ampelography of SSSR), Lyoghkaya I Pishchevaya Promishlennost: Moscow, Russia. (In Russian).

Hobosyan, S., Gasparyan, B., Harutyunyan, H., Saratikyan, A., Amirkhanyan, A., 2021: Armenian Culture of Vine and

Wine, Publishing House of the Institute of Archaeology and Ethnography: Yerevan, Armenia.

Hovhannisyan, N.A., Yesayan, A.A., Bobokhyan, A.A., Dallakyan, M.V., Hobosyan, S.G., Gasparyan, B.Z., 2017: Armenian Vine and Wine, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ): Yerevan, Armenia.

Kalinowski, S.T., Taper, M.L., Marshall, T.C., 2007: Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16, 1099-1106. DOI: 10.1111/j.1365-294X.2007.03089.x.

Maghradze, P.K., Blagonravov, P.P., 1963–1970: Ampelografiya SSSR I–III (Ampelography of SSSR I–III, Pishchepromizdat: Moscow, Russia, (In Russian).

Kumar, S., Stecher, G., Tamura, K., 2016: MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecluar Biology and Evolution 33, 1870-1874. DOI: 10.1093/molbev/msw054.

Laucou, V., Lacombe, T., Dechesne, F., Siret, R., Bruno, J.B., Dessup, M., Dessup, T., Ortigosa, P., Parra, P., Roux, C., Santoni, S., Varès, D., Péros, J.P., Boursiquot, J.P., This, P., 2011: High throughput analysis of grape genetic diversity as a tool for germplasm collection management. Theoretical and Applied Genetics 122, 1233-1245. DOI: 10.1007/s00122-010-1527-y.

Maghradze, D., Rustioni, L., Turok, J., Scienza, A., Failla, O., 2012: Caucasus and Northern Black Sea Region Ampelography. Vitis 51 (Special Issue).

Margaryan, K., Maul, E., Muradyan, Z., Hovhannisyan, A., Melyan, G., Aroutiounian, R., 2019: Evaluation of breeding potential of wild grape originating from Armenia. BIO Web Conferences 15, 01006. DOI: 10.1051/bioconf/20191501006.

Margaryan, K., Melyan, G., Röckel, F., Töpfer, R., 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.

Maul, E., Sudharma, K.N., Kecke, S., Marx, G., Müller, C., Audeguin, L., Boselli, M., Boursiquot, J.M., Bucchetti, B., Cabello, F., Carraro, R., Crespan, M., de Andrés, M.T., Eiras Dias, J., Ekhvaia, J., Gaforio, L., Gardiman, M., Grando, S., Argyropoulos, D., Jandurova, O., Kiss, E., Kontic, J., Kozma, P., Lacombe, T., Laucou, V., Legrand, D., Maghradze, D., Marinoni, D., Maletic, E., Moreira, F., Muñoz-Organero, G., Nakhutsrishvili, G., Pejic, I., Peterlunger, E., Pitsoli, D., Pespisilova, D., Preiner, D., Raimondi, S., Regner, F., Savin, G., Savvides, S., Schneider, A., Sereno C., Simon, S., Staraz, M., Zulini, L., Bacilieri, R., This, P., 2012: The European *Vitis* Database (www.eu-vitis.de) – a technical innovation through an online uploading and interactive modification system. Vitis 51, 79-86. DOI: 10.5073/vitis.2012.51.79-85.

Maul, E., Töpfer, R., Carka, F., Cornea, V., Crespan, M., Dallakyan, M., de Andrés Domínguez, T., de Lorenzis, G., Dejeu, L., Goryslavets, S., Grando, S., Hovannisyan, N., Hudcovicova, M., Hvarleva, T., Ibáñez, J., Kiss, E., Kocsis, L. Lacombe, T., Laucou, V., Maghradze, D., Maletić, E., Melyan, G., Mihaljević, M.Z. Muñoz-Organero, G., Musayev, M., Nebish, A., Popescu, C.F., Regner, F., Risovanna, V., Ruisa, S., Salimov, V., Savin, G., Schneider, A., Stajner, N., Ujmajuridze, L., Failla, O., 2015: Identification and characterization of grapevine genetic resources maintained in Eastern European Collections. Vitis 54, 5-12. DOI: 10.5073/vitis.2015.54.special-issue.5-12.

McGovern, P.E, 2003: Ancient Wine: The Search for the Origins of Viniculture. Princeton University Press. Princeton, NJ, USA.

McGovern, P., Jalabadze, M., Batiuk, S., Callahan, M.P., Smith, K.E., Hall, G.R., Kvavadze, E., Maghradze, D., Rusishvili, N., Bouby, L., Failla, O., Cola, G., Mariani, L., Boaretto, E., Bacilieri, R., This, P., Wales, N., Lordkipanidze, D., 2017: Early Neolithic wine of Georgia in the South Caucasus. Proc. Natl. Acad. Sci. USA, 114, E10309-E10318. DOI: https://doi.org/10.1073/pnas.1714728114.

Melyan, G., Safaryan, D., Nersisyan, A., 2019: Ampelography, Copyright Edition, Food and Agriculture Organization of the United Nations: Yerevan, Armenia.

Myles, S., Boyko, A.R., Owens, C.L., Brown, P.J., Grassi, F., 2011: Genetic structure and domestication history of the grape. PNAS (Proceedings of the National Academy of Sciences of the United States of America) 108, 3530-3535. DOI: 10.1073/pnas.1009363108.

Nei, M., 1973: Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA, 70, 3321–3323.

Nikoghosyan, M., Schmidt, M., Margaryan, K., Loeffl-Wirth, H., Arakelyan, A., Binder, H., 2020: SOMmelier—Intuitive Visualization of the Topology of Grapevine Genome Landscapes Using Artificial Neural Networks. Genes 11, 817. DOI: 10.3390/genes11070817.

Peakall, R.; Smouse, P.E. GenAlex 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 2006, 6, 288–295.

Poghosyan, S.H., 1962: Haykakan Sovetakan Sotsialistakan Respublikai Ampelografiya II (Ampelography of the Armenian Soviet Socialist Republic II), Haykakan Petakan Gyughatntesakan Hratararaktchutyun: Yerevan, Armenia. (In Armenian).

Poghosyan, S.H., 1981: Haykakan Sovetakan Sotsialistakan Respublikai Ampelografiya III (Ampelography of the Armenian Soviet Socialist Republic III), Hayastan: Yerevan, Armenia. (In Armenian).

Pritchard, J.K., Stephens, M., Donnelly, P., 2000: Inference of population structure using multilocus genotype data. Genetics 155 (2), 945-959. DOI: 10.1093/genetics/155.2.945.

Ramasamy, R.K., Ramasamy, S., Bindroo, B.B., Naik, V.G., 2014: STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. SpringerPlus 3, 431.

Riaz, S., De Lorenzis, G., Velasco, D., Pitcher, A., Maghradze, D., Bobokashvili, Z., Musayev, M., Goran, Z., Laucou, V., Walker, M. A., Failla, O., Preece, J., Aradhya, M., Arroyo-García, R., 2018: Genetic diversity analysis of cultivated

and wild grapevine (*Vitis vinifera* L.) accessions around the Mediterranean Basin and Central Asia. BMC Plant Biology 18, 137. DOI: 10.1186/s12870-018-1351-0.

Saitou, N., Nei, M., 1987: The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology Evolution 4, 406-425, DOI: 10.1093/oxfordjournals.molbev.a040454.

Sefc, K.M., Regner, F., Turetschek, E., Glössl, J., Steinkellner, H., 1999: Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. Genome 42, 367-373.

This, P., Lacombe, T., Thomas, M.R., 2006: Historical origins and genetic diversity of wine grapes. Trends in Genetics 22, 511-519, DOI: 10.1016/j.tig.2006.07.008.

Tumanyan, M., 1947: Haykakan Sovetakan Sotsialistakan Respublikai Ampelografiya I (Ampelography of the Armenian Soviet Socialist Republic I), Armenian SSR Academy of Sciences Press (Haykakan SSR Gitutyunneri Akademiayi Hrataraktchutyun): Yerevan, Armenia. (In Armenian).

Zohary D, Hopf M. 2000: Domestication of plants in the Old World. 3rd ed., 316pp. New York: Oxford University Press, Oxford, UK.