Vitis **54** (Special Issue), 13–16 (2015)

Molecular investigation of Caucasian and Eastern European grapevine cultivars (*V. vinifera* L.) by microsatellites

G. De Lorenzis[1)], D. Maghradze[2)], B. Biagini[1)], G. S. Di Lorenzo[1)], G. Melyan[3)], M. Musayev[4)], G. Savin[5)], V. Salimov[6)], R. Chipashvili[2)] and O. Failla[1)]

[1)] Dipartimento di Scienze Agrarie ed Ambientali, Università degli Studi di Milano, Milano, Italy
[2)] Institute of Horticulture, Viticulture and Oenology, Agricultural University of Georgia, Tbilisi, Georgia
[3)] Armenian Academy of Viticulture, Wine-making and Fruit-growing, Yerevan, Armenia
[4)] Genetic Resources Institute of the Azerbaijan National Academy of Sciences, Baku, Azerbaijan
[5)] National Institute for Viticulture and Oenology, Chisinau, Moldova
[6)] Research Institute of Viticulture and Winemaking, Baku, Azerbaijan

Summary

Grapevine (*V. vinifera* L.) is one of the most widely cultivated species of agricultural interest. The domestication of wild grapes starting in the Neolithic Age, took place in the Near East area. The aim of this study was the genotyping of germplasm coming from Armenia, Azerbaijan, Georgia and Moldova by SSR markers in order to investigate the genetic relationships among samples along the East-to-West dissemination route of grapevine after the domestication. Based on the descriptive statistics Azerbaijani samples appeared having the highest genetic diversity. PCoA and STRUCTURE analysis revealed three groups: i) Central European group; ii) a group reuniting cultivars coming from Armenia, Georgia and Moldova; iii) the group of Azerbaijani cultivars (94%). The analysis of genetic relationships in our dataset provided evidence of connection among cultivars included in the proles pontica and proles orientalis and geographical origin and human uses as well.

K e y w o r d s : domestication; genotyping; relationship; SSR.

Introduction

Grapevine (*Vitis vinifera* L.) is one of the most important widely cultivated fruit species. Archeological evidence suggests that its domestication from their wild relatives, *V. vinifera* subsp. *sylvestris*, took place in the South Caucasus about 8,000 years ago, during the Neolithic Age (This *et al*. 2006, Forni 2012). This species is highly heterozygous due to continuous breeding activities during the domestication and adaptation processes (Laucou *et al*. 2011).

The genetic diversity of the South Caucasian grapevine germplasm (wild and cultivated), considered the first centre of domestication, was investigated by molecular and ampelographic analysis (Myles *et al*. 2011, Bacilieri *et al*. 2013, Imazio *et al*. 2013, Ekhvaia and Akhalkatsi 2010). The Myles *et al*. (2011) work revealed a Near East origin of *V. vinifera* and an East-to-West dissemination gradient, following the alleged spreading routes. The results reported in Bacilieri *et al*. (2013) highlighted the genetic identity and the originality of Georgian germplasm in respect to the worldwide accessions, as well as the data described by other authors (Imazio *et al*. 2013, Ekhvaia and Akhalkatsi 2010).

The Simple Sequence Repeat (SSR) or microsatellite markers are the most frequently tools used to analyze the worldwide grapevine germplasm, as documented by different works describing the genetic characterization of the most important grapevine repositories (Laucou *et al*. 2011; Emanuelli *et al*. 2013). Microsatellites are largely used for genotyping, to solve cases of homonyms and synonyms (Laucou *et al*. 2011), to determine genetic diversity of *V. vinifera* cultivars (Cipriani *et al*. 2010, Bacilieri *et al*. 2013, Imazio *et al*. 2013) and to establish pedigree analysis (Lacombe *et al*. 2013).

In order to investigate the genetic relationships among samples along the East-to-West dissemination route of grapevine, the aim of this work was to study the genetic diversity of a set of genotypes, never studied before, coming from Armenia, Azerbaijan, Georgia and Moldova by 10 SSR markers. The data reported in this work were implemented in the frame of EU-COST project entitled "East-West collaboration for grapevine diversity exploration and mobilization of adaptive traits for breeding" (COST Action FA1003), having the aim to explore the grapevine genetic resources on a large geographic scale and in a wide range of countries (from the Caucasus to Eastern Europe to Western Europe), in order to create a network of research institutions for the identification of genetic resources, phenotypic characterization and the long-term conservation of grapevine germplasm.

Material and Methods

P l a n t m a t e r i a l s : Autochthonous grapevine cultivars coming from Armenia (29), Azerbaijan (42), Georgia (44) and Moldova (23), selected among the local germplasm of each country, were taken into account for this study. Twenty-two European varieties were included

Correspondence to: Dr. G. De Lorenzis, University of Milan, Dipartimento di Scienze Agrarie ed Ambientali, Via Celoria 2, 20133 Milano, Italy. Fax: +39-2-50316553. E-mail: gabriella.delorenzis@unimi.it

as reference. The samples were sent by cuttings to the Dip. di Scienze Agrarie e Ambientali (Milano, Italy), planted in pots and grown in greenhouse.

DNA extraction: Extraction of genomic DNA was performed per each sample using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacture's instructions. Dried young leaves (0.02 g) were ground by liquid nitrogen and the powder was used to perform the DNA extraction.

SSR amplification and detection: The samples were genotyped by 10 SSR markers: VrZag62; VrZag79; VVMD5; VVMD7; VVMD27; VVMD28; VVMD21; VVMD24; VVMD25; VVS2 (Laucou *et al*. 2011). Multiplexed PCR amplifications were performed in 25 µL final volume reaction mixture following the method described in De Lorenzis *et al*. (2012). The PCR products were carried out on ABI PRISM 310 Genetic Analyser (Applied Biosystems by Life Technologies, Foster City, USA) and the alleles were sized by GENEMAPPER 4.0 (Applied Biosystems by Life Technologies).

Data analysis: In order to estimate the genetic diversity of germplasm, the SSR data were used to determine the number of different alleles (Na), the effective number of alleles (Ne), the observed (Ho) and expected heterozygosity (He) per each provenience. These data analyses were performed by GenAlEx 6.5 software (Peakall and Smouse 2006).

The structure and the association among germplasm was investigated following two different approaches: i) Principal Coordinate Analysis (PCoA; Price *et al.* 2006), used to capture the correlation between genotypes; ii) STRUCTURE analysis (Pritchard *et al.* 2000), a Bayesian approach attempts to interpret the correlation between genotypes in terms of admixture between a defined number of ancestral populations. The PCoA analysis was carried out by GenAlEx 6.5 software, starting to the SSR correlation matrix. The STRUCTURE analysis was carried out using STRUCTURE software package (Pritchard *et al.* 2000). K value (K number of ancestral genetic groups) was chosen according to Evanno *et al.* (2005).

The genetic distance and inbreeding among the population were performed considering pairwise Nei's genetic distance and pairwise Fst analysis. The parameters were carried out in GenAlEx 6.5 software.

Results and Discussion

One hundred and sixty *V. vinifera* cultivars coming from the South Caucasus region, and Central and Eastern Europe were studied by 10 SSR loci. The allelic profiles per each locus were used to calculate descriptive statistics and the results according to the accession geographic origin are listed in Tab. 1. A total of 147 unique genotypes were detected. In each group, synonyms were identified, except for the cultivars included in the reference group. No samples showing an identical profile but different geographic origin were discovered. One hundred and sixty-six alleles and an average of 16.6 alleles per locus were detected. The number of different alleles ranged between 6.900 (Moldavian samples) and 11.200 (Azerbaijani samples), while the number of effective alleles varied between 4.493 (European varieties) and 5.948 (Azerbaijani cultivars). The mean value of Ho revealed by the analysis was high (0.769), ranging from 0.698 (Moldova) to 0.814 (Europe). The He values were very similar to Ho values, showing a mean value of 0.769 and range from 0.733 (Moldova) to 0.814 (Azerbaijan). Despite the limited number of analyzed accessions, the lower number of analyzed loci and the specific Caucasian provenience of samples, the descriptive statistics reported in this study reflected the genetic diversity highlighted in previous works, such as Laucou *et al*. (2011), where data about 2,836 SSR single profiles obtained by 20 SSR loci were descripted, Imazio *et al*. (2013), reporting results about cultivated and wild Georgian samples, and Emanuelli *et al*. (2013), regarding the molecular characterization of 1,085 accessions (*sativa*, *sylvestris*, rootstocks and hybrids genotypes).

Two different methods were performed to identify the relationship among cultivars and the structure of genetic groups. The PCoA analysis was computed on the genetic distance matrix obtained by SSR allelic profiles and the two principal coordinates of PCoA were plotted in a 2-D scattered plot (Fig. 1). The first two principal coordinates (PC) accounted for 26.21 and 24.56 % of the total variability.

The distribution differentiated the samples into three main clusters: i) Azerbaijani groups, where the most part of samples coming from Azerbaijan were grouped; ii) the rest of Caucasian samples, iii) the European (Central and

Table 1

Genetic diversity of Armenian, Azerbaijani, European, Georgian and Moldavian cultivars revealed by analysis of 10 SSR loci

Origin	No. of samples	No. of different genotypes	Na[a]	Ne[b]	Ho[c]	He[d]
Armenia	29	24	8.100	5.263	0.809	0.797
Azerbaijan	42	41	11.200	5.948	0.768	0.814
Central Europe	22	22	7.200	4.493	0.814	0.751
Georgia	44	40	8.700	4.504	0.741	0.750
Moldova	23	20	6.900	4.573	0.698	0.733
Total	160	147	8.420	4.956	0.766	0.769

[a] Number of different alleles; [b] Number of effective alleles; [c] Observed heterozygosity; [d] Expected heterozygosity.

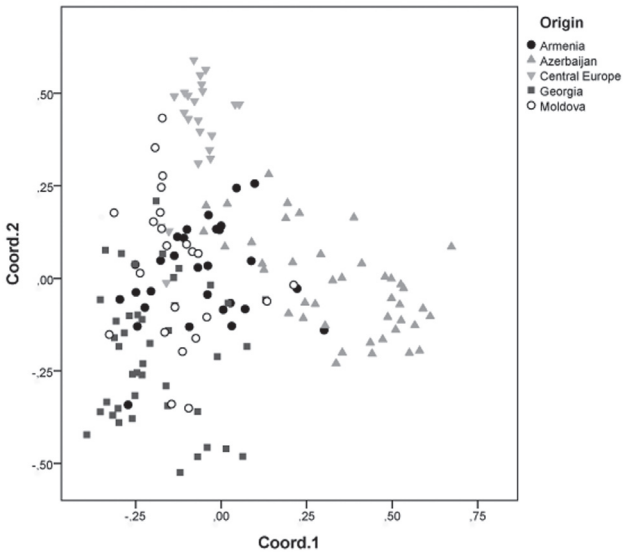

Fig. 1: Plot of the first two principal coordinates of PCoA detected by SSR data for 160 samples coming from Armenia, Azerbaijan, Georgia, Moldova and Central Europe.

Eastern) cultivars. Along the PC1, the samples coming from Azerbaijan were differentiated, while the reference cultivars appeared differently along the PC2. Moreover, the presence of overlapping zones was recognized. Several samples from Azerbaijan were genetically closed to some Eastern European (Moldova) and the Caucasian accessions, showing greater affinity. Moldova has an old tradition in grapevine cultivation and winemaking, but due to some periods of decline, depending on social, political and natural events, the most part of biodiversity was lost and further genotypes from diverse viticulture neighbouring centres were introduced (Savin 2012). The distribution of Armenian genotypes followed the geographic limits of this country. The group of samples more related to European and Caucasian varieties could be part of the gene pool from where Western accessions originated, while the group of more different genotypes could be associated to the gene pool of the enlarged center of primo-domestication, expanding to Central Asia (Bacilieri *et al*. 2013). The reference samples (Central European samples) were the samples showing the most significant differentiation in the Caucasian and East European groups. These overlapping zones disappeared removing the most two different groups (Azerbaijan and Central Europe) by the dataset (data not shown) and the Armenian, Georgian and Moldavian groups appeared distinct. The West-East gradient, following the grapevine migration from the first domestication centre, clearly identified in the analysis performed by Myles *et al*. (2011) and Imazio *et al*. (2013) was confirmed, despite clear overlapping areas slightly flattening the distribution. This differentiation of Azerbaijani samples could be due to the different usage of grapes. Indeed, the Azerbaijani cultivars are mostly table grapes, because when Azerbaijani people have started to accept the Muslim religion since 10th century AD, the population selected table grapes instead of wine grapes.

The second method used to infer the relationship among genotypes was the clustering algorithm implemented in the STRUCTURE program. Different numbers of K populations were explored (Fig. 2) and the optimal K value, evaluated following the Evanno *et al*. (2005) method, estimated the most likely number of groups at K = 3. The threshold used to estimate the group assignation was > 80 % and about 91 % of samples were assigned to a cluster at K = 3. The groups highlighted were: i) Central Europe group (92 % of the reference samples); ii) a group reuniting cultivars from different countries, containing the 73 % of Armenian samples, 96 % of Georgian genotypes and 89 % of cultivars from Moldova; iii) Azerbaijani group (94 % of cultivars). The samples showing the high admixture were the Armenian cultivars. Another STRUCTURE level was identified increasing the K value up to 5: one group per each geographical region was displayed (data not shown). In this latter case, the samples showing the high admixture level were Georgian and Armenian cultivars.

Fig. 2: Estimation of admixture proportions for 160 samples coming from Armenia, Azerbaijan, Central Europe, Georgia and Moldova by STRUCTURE analysis (K = 3).

The three-group differentiation of STRUCTURE analysis could be explained based on geographical origin and human uses: i) wine varieties from the West; ii) wine varieties from the East; iii) table varieties from the East. This latter splitting was consistent with the results reported in Bacilieri *et al*. (2013), where the analyses revealed three genetic groups, including wine cultivars from western regions, wine cultivars from the Balkans and East Europe, and table grape cultivars from the Eastern Mediterranean, Caucasus, Middle and Far East countries.

The genetic similarity among the Caucasian and European samples was also evaluated by Nei's genetic distance (Tab. 2). The Nei's genetic distance values ranged from 0.242 to 0.845, respectively for the most similar groups, Armenia and Moldova and the most dissimilar groups, Central Europe and Georgia. The Moldavian samples, geographically closer to Central Europe, showed a high genetic distance value with varieties coming from Central Europe. The dissimilarity of Azerbaijani cultivars was also confirmed by high values of Nei's genetic distance. Furthermore, the Fst values (Tab. 2) were consistent with results obtained by Nei's genetic distance. The values were very low, ranging from 0.034 (Armenia *vs.* Moldova) and 0.087 (Central Europe *vs* Georgia).

The identification of 3 different groups by the genetic analysis was in agreement with the *proles* classification proposed by Negrul (1946): i) *proles occidentalis* (cultivars from Italian Peninsula to Iberian Peninsula); ii) *proles pontica* (varieties from Georgia to Balkans and the Anatolian Peninsula); iii) *proles orientalis* (Afghanistan, Armenia, Azerbaijan and Iran).

Table 2

Pairwise population matrix of Nei's genetic distance (below the diagonal) and pairwise population matrix of Fst values (above the diagonal) obtained analyzing 5 groups of *V. vinifera* cultivars by 10 SSR loci

	Armenia	Azerbaijan	Central Europe	Georgia	Moldova
Armenia	-	0.047	0.071	0.042	0.034
Azerbaijan	0.530	-	0.072	0.065	0.056
Central Europe	0.746	0.814	-	0.087	0.075
Georgia	0.342	0.674	0.845	-	0.041
Moldova	0.242	0.505	0.629	0.281	-

Conclusions

During the last years a lot of energy was spent in order to investigate the genetic diversity, the structure and the domestication history of the whole *V. vinifera* germplasm. One limitation of these studies is that the analyzed samples, collected in repositories, were geographically limited to one or few countries and sometimes some regions, such as Minor and Central Asia, were under-represented. Hence, the results could not be considered entirely concrete about the structure of the entire grapevine gene pool. The purpose of this study was to provide additional information about the genetic diversity of European and Caucasian varieties and the relationship among these different wine-growing areas, analyzing a large set of cultivars. According to our results, a clear connection between *proles pontica* and *proles orientalis* might be assumed, as well as between geographical origin and human uses, highlighting the high genetic diversity of the South Caucasian germplasm.

Acknowledgements

Joint publication of the COST Action FA1003 "East-West Collaboration for Grapevine Diversity Exploration and Mobilization of Adaptive Traits for Breeding".

References

Bacilieri, R.; Lacombe, T.; Le Cunff, L.; Di Vecchi-Staraz, M.; Laucou, V.; Genna, B.; Péros, J. P.; This, P.; Boursiquot, J. M.; 2013: Genetic structure in cultivated grapevines is linked to geography and human selection. BMC Plant Biol. **13**, 25.

Cipriani, G.; Spadotto, A.; Jurman, I.; Di Gaspero, G.; Crespan, M.; Meneghetti, S.; Frare, E.; Vignani, R.; Cresti, M.; Morgante, M.; Pezzotti, M.; Pe, E.; Policriti, A.; Testolin, R.; 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**, 1569-1585.

De Lorenzis, G.; Imazio, S.; Biagini, B.; Failla, O.; Scienza, A.; 2012: Pedigree reconstruction of the Italian grapevine Aglianico (*Vitis vinifera* L.) from Campania. Mol. Biotechnol. **54**, 634-642.

Ekhvaia, J.; Akhalkatsi, M.; 2010: Morphological variation and relationships of Georgian populations of *Vitis vinifera* L. subsp. *sylvestris* (C. C. Gmel.) Hegi. Flora **205**, 607-617.

Emanuelli, F.; Lorenzi, S.; Grzeskowiak, L.; Catalano, V.; Stefanini, M.; Troggio, M.; Myles, S.; Martinez-Zapater, J. M.; Zyprian, E.; Moreira, F. M.; Grando, M. S.; 2013: Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. BMC Plant Biol. **13**, 39.

Evanno, G.; Regnaut, S.; Goudet, J.; 2005: Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. **14**, 2611-2620.

Forni, G.; 2012: The origin of "Old World" viticulture. In: D. Maghradze; L. Rustioni; A. Scienza; J. Turok; O. Failla (Eds.): Caucasus and Northern Black Sea Region. Vitis **Special Issue**, 27-38.

Imazio, S.; Maghradze, D.; De Lorenzis, G.; Bacilieri, R.; Laucou, V.; This, P.; Scienza, A; Failla, O.; 2013: From the cradle of grapevine domestication: molecular overview and description of Georgian grapevine (*Vitis vinifera* L.) germplasm. Tree Genet. Genomes **9**, 641-658.

Lacombe, T.; Boursiquot, J. M.; Laucou, V.; Di Vecchi-Staraz, M.; Péros, J. P.; This, P.; 2013: Large-scale parentage analysis in an extended set of grapevine cultivars (*Vitis vinifera* L.). Theor. Appl. Genet. **126**, 401-414.

Laucou, V.; Lacombe, T.; Dechesne, F.; Siret, R.; Bruno, J. P.; Dessup, M.; Dessup, T.; Ortigosa, P.; Parra, P.; Roux, C.; Santoni, S.; Vares, D.; Peros, J. P.; Boursiquot, J. M.; This, P.; 2011: High throughput analysis of grape genetic diversity as a tool for germplasm collection management. Theor. Appl. Genet. **122**, 1233-1245.

Myles, S.; Boyko, A. R.; Owens, C. L.; Brown, P. J.; Grassi, F.; Aradhya, M. K.; Prins, B.; Reynolds, A.; Chia, J. M.; Ware, D.; Bustamante, C. D; Buckler, E.S.; 2011: Genetic structure and domestication history of the grape. PNAS **108**, 3530-3535.

Negrul, A.M.; 1946: Ampelography of USSR. Pishchepromizdat, Moscow (In Russian).

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

Price, A. L.; Patterson, N. J; Plenge, R. M.; Weinblatt, M. E.; Shadick, N. A.; Reich, D.; 2006: Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. **38**, 904-909.

Salayeva, S.; Decroocq, S.; Mariette, S.; Akhundova, E.; 2010: Comparison of genetic diversity between cultivated and wild grape varieties originating from the Near-Caspian zone of Azerbaijan. J. Int. Sci. Vigne Vin. **44**, 191-200.

Savin, G.; 2012: Viticulture and winemaking in Moldova. In: D. Maghradze; L. Rustioni; A. Scienza; J. Turok; O. Failla (Eds.): Caucasus and Northern Black Sea Region. Vitis **Special Issue**, 241-272.

This, P.; Lacombe, T.; Thomas, M. R.; 2006: Historical origins and genetic diversity of wine grapes. Trends Genet. **22**, 511-519.