Large-scale screening of unknown varieties in a grapevine intra-varietal variability collection

Elsa Gonçalves[1,2], David Faísca-Silva[1], Margarida Rocheta[1], Teresa Pinto[2], Luísa Carvalho[1], Leonor Pereira[3,4], Sara Barrias[5], Sara Rocha[5], Ângela Lopes[6], Joana Guimarães[6], Teresa Mota[7], Paulo Costa[3], José Manso[2,3], Luís Arnaldo[8], Mário Cardoso[8], António Graça[9,2], César Almeida[10], Luís Vaz[10], Vanda Pedroso[10], João Brazão[6], Francisco Carlos[11], Ana Cavaco[12], Ana Chambel[12], Luís Mendes[12], Maria Godinho[12], Miguel Cachão[12], Francisco Mata[13], Cristina Fino[13], João Costa[14], José Fernandes[14], José Porfírio[14], Jorge Cunha[6], José Matos[6,15], Fernanda Simões[6], Paula Martins-Lopes[5], Wanda Viegas[1], and Antero Martins[1,2]

[1]LEAF-Linking Landscape, Environment, Agriculture and Food—Research Center, Associated Laboratory TERRA, Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal
[2]Associação Portuguesa para a Diversidade da Videira - PORVID, Lisboa, Portugal
[3]Associação para o Desenvolvimento da Viticultura Duriense – ADVID, Vila Real, Portugal
[4]CoLAB Vines&Wines, ADVID, Vila Real, Portugal
[5]DNA & RNA Sensing Lab, Polo BioISI, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal
[6]Instituto Nacional de Investigação Agrária e Veterinária, Oeiras, Portugal
[7]Associação Portuguesa de Horticultura, Lisboa, Portugal
[8]Direção Regional de Agricultura e Pescas do Norte, Braga, Portugal
[9]SOGRAPE Vinhos SA, Avintes, Portugal
[10]Direção Regional de Agricultura e Pescas do Centro, Castelo Branco, Portugal
[11]Direção Regional de Agricultura e Pescas de Lisboa e Vale do Tejo, Santarém, Portugal
[12]Associação de Viticultores do Concelho de Palmela – AVIPE, Palmela, Portugal
[13]Associação Técnica dos Viticultores do Alentejo – ATEVA, Évora, Portugal
[14]Direção Regional de Agricultura e Pescas do Algarve, Faro, Portugal
[15]CE3C- Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

Abstract. Since the last decade of the last century, it is known that many old grapevine varieties are descendants of other varieties through natural crossing. Portugal has an important program for the conservation of representative samples of intra-varietal variability of all autochthonous varieties, managed by the Portuguese Association for Grapevine Diversity (PORVID), which makes looking for genotypes with dubious identification an important activity from a perspective of its valorisation. This communication presents the results of the molecular analysis of 5,000 samples (accessions) from the PORVID's collection, using nine microsatellite *loci* currently recommended by the International Organization of Vine and Wine (OIV) for genetic grapevine identification. The results obtained confirmed the molecular identity of 4,220 samples corresponding to 214 varieties present in the official list of Portuguese varieties. In 780 samples, 95 profiles with a plural number of accessions revealed not to be listed in the Vitis International Variety Catalogue (VIVC) database, corresponding to possible varieties either descendent from natural crossing from at least one known parental variety, or from undetermined origin. Furthermore, the need for a comprehensive strategy aimed at uncovering other hidden varieties is discussed to prevent their imminent loss, deepen understanding of their origin, and add economic value and sustainability to the vine and wine sector.

1 Introduction

It is now known that many popular varieties are close relatives connected to one another by first- or second-degree relationships [1,2]. In Portugal, several research works on Portuguese autochthonous varieties have been carried out which confirmed the family relationships between many of them [3-5].

Morphological similarities between the progeny and the old parental variety are frequently observed, being in some cases difficult to distinguish between them in the field, even by trained ampelographers. A typical example of this situation, among many others, is that of the Touriga Fêmea variety, resulting from the crossing of

Touriga Nacional × Marufo, which is often confused in the field with Touriga Nacional. The knowledge already acquired about the occurrence of natural crossings and cases of difficult distinction between parents and progeny authorizes the hypothesis of the existence of more yet unidentified progenies (varieties) mixed with their parents in vineyards.

Identifying these suspicious plants in vineyards for further molecular diagnosis is challenging, due to the large environmental deviations that modify the phenotype of individual plants and the scale of the required experimental work (implying necessarily the strict and repeated observation of many thousands of plants distributed across the country).

The solution for recovering genetic resources in old vineyards is to prospect and conserve *ex-situ* representative samples of the intra-varietal variability of all autochthonous varieties. Multiple plants of each genotype, preserved in pots and in field trials, are less affected by environmental deviations and more likely to reveal suspected identification that can be marked for molecular diagnosis. Fortunately, Portugal has a large and diverse pool of autochthonous varieties (approximately 250) which has been subjected to an important conservation program of representative samples of intra-varietal variability, managed by the Portuguese Association for Grapevine Diversity (PORVID). Presently, more than 30,000 accessions of over 218 identified varieties are already conserved in pots and in field trials [6], which justifies the varietal identification of dubious accessions. In fact, morphological annotations of thousands of genotypes are already available and many of those are conserved in field trials for morphological, cultural, and oenological traits evaluation with high discriminating power, thus revealing the presence of possible new varietal identities.

The objective of this work is to demonstrate the importance of the large-scale prospection of intra-varietal variability of autochthonous varieties in old vineyards, and its respective conservation, aiming to save important unrevealed genetic resources, namely uncovering hidden varieties mixed with their parents in old vineyards. This strategy prevents the imminent loss of such germplasm, deepening the understanding of their origin, and adding economic value and sustainability to the vine and wine sector.

2 Materials and methods

2.1 Plant material

The genotypes tested were collected in the PORVID's grapevine intra-varietal variability collection composed of more than 30,000 accessions conserved in pots and/or in field trials. All those accessions were prospected in a large number of old vineyards that were planted before selection and nursery activities (because only those vineyards preserve the diversity that was created in the past), following an appropriate methodology of prospection of intra-varietal variability [7]. Prospection was performed by a national network composed of more than 120 technicians/ampelographs and was conducted in wine-demarked regions of Portugal (Alentejo, Algarve, Bairrada, Beira Interior, Dão, Douro, Lafões, Lisboa, Península de Setúbal, Douro, Tejo, Trás-os-Montes, and Vinhos Verdes).

For this study, 5,000 accessions were selected for molecular analysis according to the following criteria: (1) genotypes whose varietal identification was not achieved by technicians/ampelographs; (2) genotypes whose varietal identification raised doubts at the time of prospection; (3) genotypes in large field trials for selection, which proved to be atypical in relation to the varietal pattern based on morphological observations and on evaluations of cultural and oenological traits.

Samples of young leaves from those 5,000 accessions were collected between May and June 2020 and 2021, and stored at -80ºC.

2.2 Molecular analysis

The strategy used to screen such a large number of accessions was an expedited and extensively applied methodology in grapevine identification, the microsatellite markers (SSRs).

Leaves from each sample (genotype/accession) were macerated in liquid nitrogen and total genomic DNA was extracted and purified with DNeasy Plant Mini Kit (Qiagen), following the manufacturer's instructions. Nucleic acid concentration was measured using a microplate reader Synergy HT (Biotek, Germany), with the software Gen5™ (Biotek, Germany) and integrity was accessed in an agarose gel 0.7% (p/v). DNA was stored at 4ºC.

The SSRs were selected following the OIV recommendation for genetic grapevine identification [8]: VVS2, VVMD5, VVMD7, VVMD25, VVMD27, VVMD28, VVMD32, VrZAG62, and VrZAG79. The amplification was carried out using multiplex PCR. Each forward primer was labelled with a fluorochrome: 6 - FAM (VVMD5, VVMD25, and VVMD32), HEX (VVMD7, VVMD28, and VrZAG79), and Atto 550 (VVMD27, VVS2, and VrZAG62). Multiplex PCR was used in combinations of three by three, according to the expected amplification size: reaction 1 (VVMD5, VVMD7, VVS2, and VrZAG79), reaction 2, and reaction 3 (VVMD25, VVMD27, VVMD28).

The PCR reaction was composed by 10 ng of DNA, 10 µL PCR Master Mix (Qiagen), 0.5 µL of each primer (10 µM), and 7 µL of RNA-free water. All amplifications were carried out using a thermocycler T100 (BioRad) in a 96-well plate under the following conditions: an initial step of 95ºC for 15 minutes, followed by 34 cycles of 94ºC for 30 seconds, 57ºC for 90 seconds, and 72ºC for 60 seconds, with a final extension step of 72ºC for 30 minutes.

Fragment analysis was carried out in a ABI 3730XL sequencer (Applied Biosystems), after adding 10-15 µL formamide to each sample. ABI ROX-500 was the molecular size marker used. The fragment analysis data were retrieved in .fsa files and analysed with the OSIRIS software (https://www.ncbi.nlm.nih.gov/osiris/). Data were processed for each sample and alleles were scored. The genetic profile of each sample was based on the

peaks presented in the electropherogram for each nSSR marker. After allele scoring, the microsatellite profiles of samples were adjusted by comparing the genetic profiles of the control grapevine varieties (Pinot Noir, Chardonnay, Muscat a Petits Grains Blanc, and Cabernet Sauvignon) with their respective reference profile in the Vitis International Variety Catalogue database (VIVC). After standardization, the SSR profile of each sample was screened against the nSSR profiles available in the VIVC nSSR database [9].

Parentage assignment was conducted by the CERVUS software [10,11] (http://www.fieldgenetics.com) with the aim of identifying possible first-order kinship relationships: trios (mother-father offspring) and duos (parent-offspring pairs). A total of 100,000 computer simulations were used to determine the critical values of LOD score for strict (95%) and relaxed (85%) confidence levels. A maximum of three nSSR *loci* mismatches was allowed for duos and trios.

3 Results and Discussion

The results obtained allowed the clarification of the molecular identity of 4,220 samples which correspond to 214 varieties present in the official list of Portuguese varieties. Since varietal identity was confirmed, two immediate consequences arise from these results, which value the intra-varietal variability conservation collection: (1) the selection process of several grapevine varieties that are not under selection can be pursued, based on a representative set of identified accessions of the intra-varietal variability of the variety; (2) the distribution of well-identified propagation material of minority varieties, not selected, but containing intra-varietal diversity and, consequently, environmental stability (less sensitive to genotype × environment interaction).

In 780 samples, 95 new profiles with a plural number of accessions, not listed in the VIVC database, were identified. Interestingly, the same profile was sometimes observed in samples from different old vineyards located in different locations, including different wine regions of Portugal. Some examples of this finding are provided in Table 1. The plural number of accessions with the same profile in different locations reinforces the idea that the genetic material corresponds to an ancient variety that was in cultivation in the past with large geographical coverage.

In Table 2, some new profiles not listed in the VIVC database are illustrated, as well as the result of the study of kinship assignment. When performing this analysis two situations were observed: (1) the possibility of having accessions either descendent from natural crossing from at least one known parent; (2) accessions from undetermined origin.

This first result confirms what really underpinned the construction of this experimental work: the perception of the existence of autochthonous varieties that are progenies of the natural crossing of other varieties and that until now have been confused with their parents in the old vineyards. Typical examples are the results obtained for Code Group Gtfe, which corresponds to an accession that is in the intra-varietal variability collection of Touriga Fêmea variety, or for Code Group Gnm, which is an accession present in the intra-varietal collection of Negra Mole variety. In both cases, accessions have as one of their possible parents the ancient variety by which they were misidentified in the field. Additionally, many other examples of this type can be described. For example, in accessions whose varietal identification raised doubts at the time of prospection as Arinto do Interior, Síria, Bastardo, Marufo, Encruzado varieties (cases of Code Groups G7, G22, G9, G24, and G16, respectively), the results confirmed that they do not match these varieties but one of their possible parents is the ancient variety by which they were confounded. This outcome leads to the perspective of the existence of more accessions of identical nature that should be repeatedly sought through information gathering and prospecting strategies throughout the country.

Table 1. Examples of cases where the same profile was found for different samples (accessions), respective wine-growing regions and number of different vineyards where the accessions were prospected.

Code Group	No. Samples	Origin Regions (No. different vineyards in the region)
G7	16	Dão (8), Douro (1)
G25	14	Vinhos Verdes (5)
G5	10	Alentejo (6), Lisboa (1)
G22	9	Dão (6), Beira Interior (1), Lisboa (1)
G6	9	Beira Interior (5)
G23	8	Dão (5)
G15	8	Dão (6), Algarve (1)
G14	7	Beira Interior (5) and Dão (1)
G8	7	Dão (5)
G1	6	Lisboa (5)
G3	6	Douro (4) and Vinhos Verdes (1)
G9	6	Beira Interior (3)
G4	5	Dão (3)
G32	5	Vinhos Verdes (2)
G20	5	Beira Interior (2)
G30	4	Algarve (3), Douro (1)
G24	4	Beira Interior (1)
G16	3	Dão (3)
G21	2	Tejo (1), Vinhos Verdes (1)
G2	2	Alentejo (2)

The second result is composed of profiles found in samples that do not correspond to progenies of crosses between varieties listed in VIVC (example of Code Groups G15, G20, G23, and G32, listed in Table 2). This new group of genotypes may therefore have resulted from multiple crossings between other rare varieties not included in official lists, some possibly already lost, which will make the reconstitution of the respective phylogenetic origin more difficult. However, these

genotypes are highly enriching contributors to the gene pool of the old varieties, and probably autochthonous, and therefore should not be disregarded. In fact, analysing these examples, it is important to highlight that these accessions were found in different vineyards, and, in some cases, in different wine-growing regions. Once again, this finding supports the perspective of the existence of more identical accessions that must be sought in other old vineyards.

The results showed in this text are only a small part of the total work performed and exemplify the potential of large-scale screening of intra-varietal variability. Some analyses will be checked with a larger set of SSRs and different molecular markers trying to clarify a more accurate kinship assignment.

As a general conclusion, it is important to highlight that the results of this work demonstrate an increase in the range of varieties in the country assigned local or Iberian and, consequently, also the existence of diversity suitable to be applied for the selection of relevant traits and in the adaptation of the grapevine varieties to the demands of present and future challenges.

This research was supported through funding of the projects "PRODER PA Nº 020000018572, PA Nº 020000018999 and PA Nº 020000018629, "Participative Budget Portugal, Project 38 - In Lisbon and in the Tejo Valley, new grapevine varieties for new wines - In demand of the secrets of the natural evolution of the Portuguese vine", "Conservation and selection of ancient grapevine varieties" (PDR2020-784-042704), and "Save the intra-varietal diversity of autochthonous grapevine varieties (PRR-C005-i03-|-000016).

References

1. S. Myles, A. Boyko, C. Owens, P. Brown, F. Grassi, B. Aradhya, A. Reynolds, J. Chia, D. Ware, C. Bustamante, E. Buckler, PNAS **108**, 3530-3535 (2011)
2. T. Lacombe, J. Boursiquot, V. Laucou, M. Di Vecchi-Staraz, J.P. Péros, P. This, Theor. Appl. Genet. **126**, 401-414 (2013)
3. I. Castro, J.P. Martín, J.M. Ortiz, O. Pinto-Carnide, Sci. Hortic. **127**, 507-514 (2011)
4. J. Cunha, J. Ibáñez, M. Teixeira-Santos, J. Brazão, P. Fevereiro, J.M. Martínez-Zapater, J.E. Eiras-Dias, Aust. J. Grape Wine Res. **22**, 504-516 (2016)
5. S. Barrias, L. Pereira, S. Rocha, T. A. Sousa, J. Ibáñez, P. Martins-Lopes, Sci. Hortic. **311**, 111826 (2023)
6. E. Gonçalves, A. Martins, Plants **11**, 1917 (2022)
7. A. Martins, E. Gonçalves, Grapevine breeding programmes in Portugal. In *Grapevine Breeding Programs for the Wine Industry*: Traditional and Molecular Techniques. A. G. Reynolds ed., Woodhead Publishing Elsevier, UK, 159-182 (2015)
8. OIV General Assembly, *OIV Protocol for Identification of Varieties*, Geneva, Switzerland (2019)
9. E. Maul, R. Topfer, *Vitis International Variety Catalogue* (VIVC) www.vivc.de (accessed 19.01.23) (2023)
10. T.C. Marshall, J. Slate, L.E.B. Kruuk, J.M. Pemberton, Mol. Ecol. **7**, 639-655 (1998)
11. S.T. Kalinowski, M.L. Taper, T.C. Marshall, Mol. Ecol. **16**, 1099-1106 (2007)

Table 2. Examples of some new profiles not listed in the VIVC database and possible kinships (trios or duos) using the 9 SSRs recommended by the OIV.

Code	VVS2		VVMD5		VVMD7		VVMD25		VVMD27		VVMD28		VVMD32		VrZAG62		VrZAG79		Parent 1 × Parent 2
G7	151	151	224	240	239	257	255	263	180	184	234	248	240	272	188	194	247	251	Arinto do Interior × ?
G25	133	151	228	240	239	257	239	249	186	190	236	258	240	240	188	194	251	251	Arvine Petit × ?
G5	137	145	240	242	239	243	241	241	182	186	234	234	240	272	188	196	237	247	Corazon de Cabrito × ?
G22	143	151	224	240	239	253	249	263	182	190	228	248	240	252	186	200	247	251	Síria × ?
G6	143	151	238	242	239	239	249	255	186	190	258	258	256	260	186	188	245	251	Castellana Blanca × ?
G23	151	151	234	234	239	243	249	249	180	182	244	244	240	252	188	196	247	251	? × ?
G15	151	151	224	228	239	253	249	263	182	186	248	258	240	272	188	200	247	259	? × ?
G14	143	157	228	242	239	257	249	255	180	180	248	258	256	272	186	188	251	251	Folha de Figueira × ?
G8	145	151	228	242	239	257	255	263	180	184	234	234	240	256	186	188	247	251	Fonte Cal × ?
G1	143	151	228	228	253	257	249	255	180	182	234	248	252	272	188	200	251	259	Airen x Alfrocheiro
G3	133	145	230	242	239	249	241	241	186	190	234	248	256	272	188	194	243	251	Listrão × ?
G9	133	151	234	240	239	257	249	255	176	186	234	254	240	256	186	188	243	247	Bastardo × ?
G4	133	151	238	240	239	243	241	249	182	186	234	236	256	272	188	188	237	247	Corazon de Cabrito × ?
G32	143	151	234	242	239	239	249	249	176	182	246	248	256	272	188	196	243	247	? × ?
G20	133	151	228	242	239	243	239	249	186	190	258	258	256	272	188	194	245	259	? × ?
G30	133	147	230	240	239	239	241	255	190	195	248	252	252	260	192	196	243	257	Larião × ?
G24	133	145	230	238	243	257	239	241	184	190	248	254	252	256	192	194	245	247	Marufo × ?
G16	151	151	228	234	239	257	241	249	184	190	234	254	240	272	194	194	247	251	Encruzado × ?
G21	133	151	224	228	239	263	239	263	184	186	228	234	240	252	194	196	247	251	Espadeiro × ?
G2	137	151	224	236	239	243	255	263	182	184	228	234	240	252	186	188	247	247	Arinto do Interior × ?
Gba	143	151	228	228	239	239	241	263	184	195	228	236	240	256	188	194	247	247	Fernão Pires × Branda
Gtfe	143	143	238	238	239	257	255	255	180	190	234	268	240	256	188	194	245	251	Touriga Fêmea × Touriga Nacional
Gnm	143	145	224	224	239	239	241	255	182	182	244	258	252	272	188	196	247	259	Manteúdo × Negra Mole