CENTENARY OF THE OIV
45th WORLD CONGRESS OF VINE AND WINE
14-18 OCTOBER 2024 FRANCE-DIJON
OIV
Available on IVES Conference Series vine & wine
DOI : https://doi.org/10.58233/4yRbJwf5

Plastid genomics of *vitis vinifera* L. for the understanding the molecular basis of grapevine domestication

Ia Pipia[1*], Thierry Lacombe[2], Ana Gujabidze[1], Tengiz Nozadze[1] and Vazha Tabidze[1]

[1] Institute of Molecular Genetics, Agricultural University of Georgia, 0159, Tbilisi, Georgia
[2] AGAP, Université de Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

Abstract. Over the past decade, numerous scientific studies have utilized advanced next-generation plastid DNA technologies to assess the genetic diversity of grapevine (*Vitis vinifera* L.) worldwide, aiming for a deeper understanding the possible molecular mechanisms of domestication process. For such studies, it is essential to include both cultivated varieties (*Vitis vinifera* subsp. *vinifera*) and their wild ancestors (*Vitis vinifera* subsp. *sylvestris*) from all potential distribution and domestication sites, as only by comparing the genetic profiles of cultivated and wild specimens can the potential sites of domestication be identified. The main goal of the presented research was the next-generation sequencing (NGS) of complete plastid genomes mainly from European and Mediterranean wild grapevines, as well as cultivated varieties from Europe and Georgia (South Caucasus), and American and Asian *Vitis* species, followed by *in silico* genome assembly and phylogenetic analyses. The results revealed that plastomes of European and Mediterranean wild grapevines predominantly exhibit the so-called GTA haplotype; Portuguese wild grapevines exhibit a new type of ATA plastid haplotype; Plastomes of Asian and American species are distinct from those of European and Mediterranean grapevines and the Rkatsiteli haplotype (AAA) remains genetically unique among the plastomes sequenced to date.

1. Introduction

The molecular basis of the grapevine domestication process remains incompletely understood. For the geographic origin of grapevine domestication and beginnings of winemaking, the South Caucasus, eastern Taurus (in present-day Turkey) and the Zagros Mountains (in modern-day Iran) are considered, along with evidence supporting multiple origins of cultivated grapevines, with one in the Near East and another in the Western Mediterranean region [1-4].

In the past decade, numerous scientific studies have utilized advanced next-generation plastid DNA technologies to assess the genetic diversity of a global set of cultivated and wild grapevines [5-12]. Plant's plastid genomes, which follow maternal unisexual inheritance, are is much smaller than the nuclear genome, and their genomic organization and structure are relatively conserved. An important feature of plastoms is their low rate of point mutations. All these characteristics are make plastoms extensively used for accurate studies of systematic phylogenetic lineages [13]. The main findings of the aforementioned studies can be summarized as follows: worldwide set of cultivated (*Vitis vinifera* subsp. *vinifera*) and wild grapevines (*Vitis vinifera* subsp. *sylvestris*) share four identical plastid haplotypes, which are defined by single nucleotide polymorphisms (SNPs) located in non-coding plastid DNA regions (the *trn*H-*psb*A intergenic spacer, the *rpl*16 intron, the *acc*D-*psa*I intergenic spacer). Each haplotype contains a number of grapevine samples and is named after the most well-known cultivar within that haplotype: Chkhaveri-Pinot Noir haplotype (GTA), Meskhuri Mtsvane-Chardonnay haplotype (ATA), Saperavi-Cabernet Sauvignon haplotype (ATT), and Rkatsiteli haplotype (AAA) [5, 8]. All above mentioned haplotypes were detected in Georgian (South Caucasus) cultivars of *V. vinifera* subsp. *vinifera* and only three of them (GTA, ATA, ATT) in worldwide cultivars, including Georgian ones. Remarkably, some coincidences were found between the chlorotypes identified by Arroyo-García *et al*., and the haplotypes GTA, ATA, and ATT. Nothing can be concluded about the possible overlap of the Rkatsiteli haplotype with any chlorotype due to the absence of Georgian grapevine samples in the chlorotype studies [4, 12]. The dominance of the GTA haplotype was observed in most wild grapevine samples from Europe and Mediterranean basin (i. e. France, Germany, Italy, Spain, Portugal, Morocco, Tunisia, Algeria, Turkey). Only the ATA haplotype was discovered in Greece [8, 11-13]. These findings are significant because identifying a potential domestication center for grapevines requires evidence of genetic diversity in a specific geographic region and shared haplotypes between wild and cultivated varieties. The Rkatsiteli haplotype (AAA) deserves special mention. This haplotype was found only in Georgian cultivated and South Caucasian wild grapevines what makes it a genetically unique haplotype with no analogue in the world [8, 11-13].

Our previous researches (in which numerous plastomes were subjected to Next generation sequencing (NGS), with most deposited in NCBI) indicate that to achieve a more refined understanding of the genetic diversity and linkages of *Vitis* plastomes, and to uncover potential molecular mechanisms of domestication, it is essential to analyze as many plastomes as possible. To this end, for a deeper look into the plastid genome diversity of *Vitis*, the presented research involved next-generation sequencing (NGS) of complete plastid genomes mainly from European and Mediterranean wild grapevines, as well as cultivated varieties from Europe and Georgia (South Caucasus), and American and Asian *Vitis* species, followed by *in silico* genome assembly and phylogenetic analyses.

2. Materials and methods

Cuttings, dried and fresh leaves of the *Vitis* samples used in this research were obtained from the Grapevine Collection of Vassal-Montpellier at the National Research Institute for Agriculture, Food and Environment (INRAE), France, the National Clonal Germplasm Repository at the University of California, Davis, USA and National Centre for Grapevine and Fruit Tree Planting Material Propagation, Georgia. Genomic DNA extractions were performed by using the CTAB-based protocol [15]. Next-generation sequencing, including the construction of shotgun genomic DNA libraries and sequencing on an Illumina HiSeq, was conducted at the Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign (UIUC), USA. For the assembly of the plastid genomes, the SOAPdenovo genome assembler (http://soap.genomics.org.cn/soapdenovo.html) was used. Each assembled plastom was referenced against the genome of the same haplotype (e.g., for GTA haplotype plastom of Maxxa, GenBank DQ424856.1; for ATA haplotype plastom of 'Meskhuri Mtsvane', GenBank AB856291; for ATT haplotype plastome of 'Saperavi', GenBank AB856290.1; for AAA haplotype plastome of 'Rkatsiteli', GenBank AB856289.1). For the phylogenetic analyzes different approaches of comparative genomics were used (i.e. Mafft, Blast).

3. Results and discussion

Table 1 presents the detailed list of the analyzed grapevine plastomes sequenced in the framework of the presented work.

Table 1. List of all analyzed grapevine plastoms for this study.

Sample		Country, region	Sex
Vitis vinifera subsp. *sylvestris*	V. silvestris Montagnella 4 (8500Mtp438)	Italy	Male
	V. sylvestris Valletta Drago (8500Mtp444)	Italy	Male
	Lambrusque Delta Stabiaccie E (8500Mtp512)	France, Corsica	Male
	V. sylvestris Portugal 110104 (T) (8500Mtp412)	Portugal	Male
	V. sylvestris Portugal 110504 (3) (8500Mtp425)	Portugal	Male
	V. sylvestris Portugal 110505 (3) (8500Mtp426)	Portugal	Female
	Lambrusque Grésigne 6 (8500Mtp126)	France	Male
	Vitis silvestris Gmelin (8500Mtp41)	Germany	Female
	Lambrusque Ul'any nad Zitavou A38 (8500Mtp200)	Slovenia	Male
	Lambrusque Ul'any nad Zitavou A46 (8500Mtp205)	Slovenia	Male
	Lambrusque de semis Tekkale Kilisi 13-5 (23429Mtp13-5)	Turkey	-
Vitis vinifera subsp. *vinifera*	Muscat d' Alexandria	Egypt	-
	Muscat Petite Blanc	Greece	
	Ojaleshi	Georgia	-
	Tavkveri	Georgia	-
	Aladasturi	Georgia	-
	Tsitska	Georgia	-
American & Asian species	*V. romaneti*	Asia	-
	V. thumbergii	Asia	-
	V. californica	America	-

3.1. Haplotype affiliations

At the first step of the research the complete chloroplast genome reads of all analyzed plastoms were subject of *in silico* genome assembly by SOAPdenovo genome assembler. It was determined that the each analyzed chloroplast genome is mainly 160.928 bp long. In some genomes, the length varied by up to ± 50 bp due to observed SNPs and InDels (not discussed in this paper). After the *in silico* genome assembly, haplotype definitions for studied plastomes were determined. The haplotypes were defined based on SNPs found in the *trn*H-*psb*A intergenic spacer at position 205 bp and at two positions, 86715 and 86721 bp, within the *rpl*16 intron of the chloroplast genome. In the three-letter name of each haplotype, the first letter corresponds to the SNP found in the *trn*H-*psb*A genome region and the remaining two letters correspond to SNPs in the *rpl*16 intron region. Table 2 presents distribution of haplotypes among all studied grapevine chloroplast genomes and the approximate age of plastid haplotypes adopted from the paper of Zecca *et al*., [7].

It had been shown that in European and Mediterranean wild grapevines (*Vitis vinifera* subsp. *sylvestris*), the GTA or Chkhaveri-Pinot Noir haplotype is prevalent. This haplotype is also found in two Georgian cultivars - 'Aladasturi' and 'Tsitska' ('Cicka'). The other two haplotypes, ATA and ATT, appearing only in wild grapevines from Turkey and Corsica (France), also in varieties 'Muscat d'Alexandria' and 'Muscat Petite Blanc' (*Vitis vinifera* subsp. *vinifera*). Haplotype AAA or Rkatsiteli haplotype is characteristic only of Georgian cultivated varieties 'Ojaleshi' and 'Tavkveri'. It should be emphasized that the haplotype detected in 'Muscat d'Alexandria,' 'Muscat Petite Blanc,' 'Ojaleshi,' 'Tavkveri,' 'Aladasturi,' and 'Tsitska' was previously identified through sequencing of only the aforementioned non-coding plastome regions [5]. In this study, these plastomes were completely sequenced, confirming the haplotypes of the mentioned varieties and providing information about similarities throughout the complete chloroplast genomes.

The described results align with our previous works, showing the GTA haplotype is dominant in European and Mediterranean wild grapevines; cultivated and wild grapevines globally share plastid haplotypes; and the Rkatsiteli haplotype is genetically unique among sequenced plastomes [5, 8-13].

Table 2. Distribution of haplotypes among all studied grapevine chloroplast genomes (*Vitis vinifera* subsp. *sylvestris*, *Vitis vinifera* subsp. *vinifera*) along with the approximate age of plastid haplotypes. An asterisk (*) indicates that the haplotype matches those reported in our earlier studies [5, 7, 11, 12].

Geographic region	Country	Plastid haplotype	The approximate age of plastid haplotypes
European and Mediterranean wild grapevines	Italy *	GTA	5-4 Ma
	Portugal *		
	France *		
	Germany *		
	Slovenia		
Georgian varieties 'Aladasturi' and 'Tsitska'	Georgia		
Mediterranean wild grapevines	Portugal	ATA	3-1 Ma
	Turkey *		
Variety 'Muscat d'Alexandria'	Egypt *		
Mediterranean wild grapevines	France, Corsica	ATT	
Variety 'Muscat Petite Blanc'	Greece *		
Georgian varieties 'Ojaleshi' and 'Tavkveri'	Georgia *	AAA	

The presence of all four haplotypes (GTA, ATA, ATT, AAA) simultaneously in the same geographic location, as demonstrated by our previous and current works, has been reported only for Georgia and the South Caucasus, in both cultivated varieties and wild grapevines to date [8-13]. Haplotype diversity is a crucial factor in the development of new forms and the domestication of plant species. Based on the experimental data provided above, the South Caucasus is principally noteworthy in this regard. According to the analyses conducted in the scope of this work, some biodiversity of chloroplast genomes was also detected in Portuguese wild grapevines, as they display the presence of two wild plastid haplotypes - GTA and ATA. Interestingly, the ATA haplotype detected in two samples of Portuguese wild grapevine (*V. sylvestris* Portugal 110504 (3) (8500Mtp425), *V. sylvestris* Portugal 110505 (3) (8500Mtp426)) could represent a new and potentially older variant of the ATA haplotype. This hypothesis is supported by the fact that the newly detected ATA haplotype features a 33 bp length duplication but lacks the 54 bp length deletion commonly associated with the ATA haplotype. We suggest that this new variant may represent an ancestral form of the ATA haplotype, from which the general ATA haplotype evolved through the acquisition of the 54 bp length deletion (Fig. 1).

Based on SNPs at three non-coding regions of the chloroplast genome, Asian and American species of *Vitis* might be attributed to the ATA haplotype. However, this attribution may not be accurate, as the plastomes of these species do not match other characteristics (especially InDels) of complete chloroplast genomes associated with ATA haplotype.

3.2. Phylogenetic analyzes

To demonstrate the possible genetic linkage between the plastomes of European, Mediterranean and Georgian wild and cultivated grapevines, a phylogenetic tree was constructed by using the Neighbor-Joining method [16]. On the tree of Fig. 2, two main clades can be distinguished. One clade includes plastomes from European and Mediterranean wild grapevines and Georgian cultivars ('Aladasturi', 'Tsitska'), all representing the GTA haplotype. The second clade contains sub-clades of ATA, ATT and AAA haplotypes and is primarily composed of plastomes of European and Mediterranean wild grapevines, cultivated varieties ('Ojaleshi', 'Tavkveri', 'Aladasturi', 'Tsitska', 'Saperavi', 'Meskhuri Mtsvane', 'Rkatsiteli' – all Georgian) and 'Muscat d'Alexandria'.

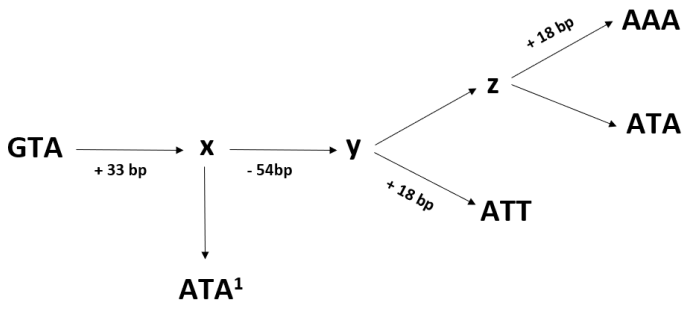

Fig. 1. Phylogenetic scheme based on single nucleotide polymorphisms (SNPs) and InDels showing the order of grapevine chloroplast DNA haplogroups. The ATA¹ haplotype, representing Portuguese wild grapevines, lacks the 54 bp deletion characteristic of the general ATA haplotype. Hypothetical ancestral haplotypes are designated as x, y, and z.

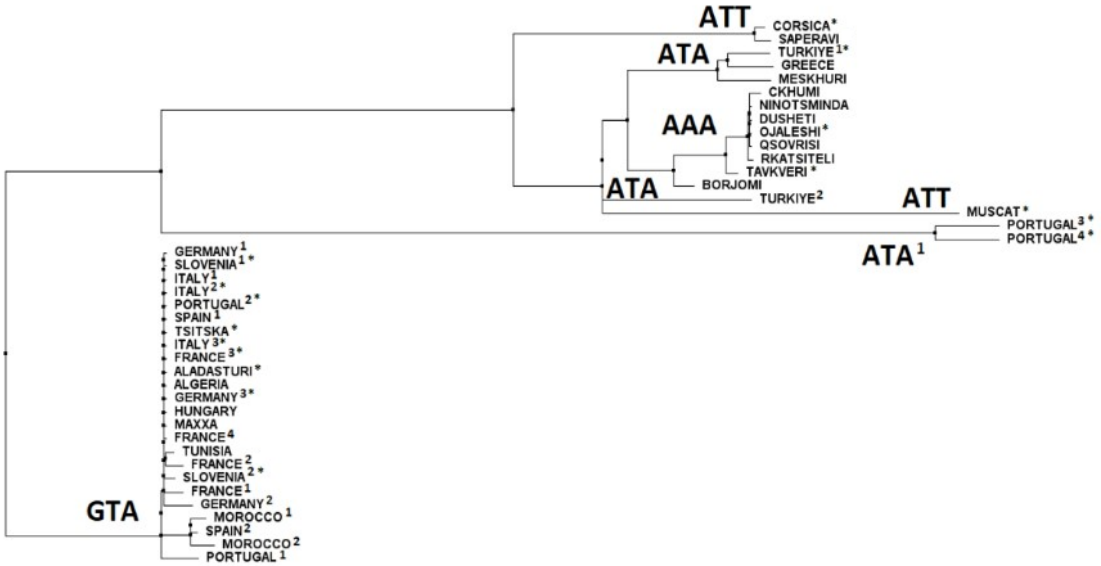

Fig. 2. Complete chloroplast genome phylogeny of genus *Vitis* constructed by Neighbor Joining tree using Jalview version 2. An asterisk (*) indicates the plastoms of the presented work, other ones are taken from our earlier studies [9-13]. Plastoms of wild grapevines are mention by country of origin. Superscripts show the sample counts for each specific country. 'Meskhuri' refers to the Georgian cultivar 'Meskhuri Mtsvane,' while 'Muscat' refers to 'Muscat d'Alexandria'.

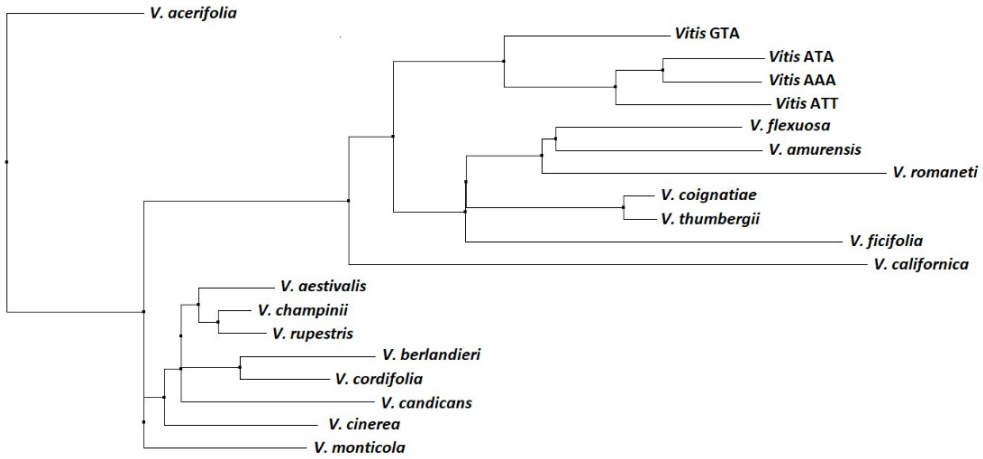

Fig. 3. Phylogeny of general plastid haplotypes of *Vitis vinifera* L. and Asian and American species. The tree was constructed by Neighbor Joining tree using Jalview version 2.

The genetic linkage between general plastid haplotypes of *Vitis vinifera* L. and Asian and American species is presented in Fig. 3. The phylogenetic tree was constructed by using the Neighbor-Joining method and include materials of the presented work (*V. romaneti*, *V. thumbergii*, *V. californica*) as well as our previous research [7, 16]. Signs '*Vitis* GTA', '*Vitis* ATA', '*Vitis* ATT' and '*Vitis* AAA' correspond to the general plastid haplotypes, which were delivered to the program as one general plastome from each haplotype (for GTA – Maxxa [GenBank# DQ424856.1], for ATA Meskhuri Mtsvane [GenBank# AB856291], for ATT Saperavi [GenBank# AB856290.1], and for AAA Rkatsiteli [GenBank# AB856289.1]). It had been shown that all four general haplotypes (GTA, ATA, ATA, AAA) form an isolated clade, with the GTA haplotype as the ancestral one. Close to this clade is the group of Asian species, *V. flexuosa*, *V. amurensis*, *V. romaneti*, *V. coignetiae, V. thumbergii, V. ficifolia*. American species make separated from the mentioned two clades group, including *V. aestivalis, V. champinii, V. rupestris, V. berlandieri, V. cordifolia, V. candicans*, *V. cinerea*, *V. monticola*. Surprisingly, that American species *V. californica* is positioned on the tree not near the American specimens but close to the Asian ones. One additional American species, *V. acerifolia*, appears to have a distinct plastome with no significant genomic similarities to other plastomes of American *Vitis*, as it is positioned separately on the phylogenetic tree.

4. Conclusions

To summarize the data from NGS and comparative genomics studies of grapevine plastomes, the following conclusions can be made:

- Plastomes of European and Mediterranean wild grapevines predominantly show the presence of GTA haplotype;
- Portuguese wild grapevines exhibit a new type of ATA haplotype alongside the GTA haplotype;
- Plastomes of Asian and American species are quite distinct from general plastid haplotypes. At the same time Asian species show a closer genetic distance to the plastid haplotypes of *Vitis vinifera* L.;
- The Rkatsiteli haplotype (AAA) remains genetically unique among the plastomes sequenced to date, as it has been found only in Georgian and South Caucasian wild grapevines and cultivated varieties.

5. Acknowledgements

Authors kindly thank Shota Rustaveli National Science Foundation of Georgia (SRNSFG) [grant number: FR 19-14378] for supporting this research.

6. References

1. P. E. McGovern, *Ancient Wine* (NJ: University Press, 2003)
2. H. Barnard, A. N. Dooley, G. Areshian, B. Gasparyan & K. F. Faull, *J. Archaeol. Sci*., **38** (2011)
3. S. Myles, A. R. Boyko, Ch. L., Owens, P. J. Brown, F. Grassi, M. K. Aradhya, B. Prins, A. Reynolds, K. Chia, D. Ware, C. D. Bustamante adna E. S. Buckler, *PNAS*, **108(9)** (2011), https://doi.org/10.1073/pnas.1009363108
4. R. Arroyo-García, L. Ruiz-García, L. Bolling, R. Ocete, M. A. López, C. Arnold, A. Ergul, G. Söylemezoğlu, H. I. Uzun, F. Cabello, J. Ibáñez, M.K. Aradhy,, A. Atanassov, I. Atanassov, S. Balint, J. L. Cenis, L. Costantini, S. Goris-Lavets, M. S. Grando, B.Y. Klein, P. E. McGovern, D. Merdinoglu, I. Pejic, F. Pelsy, N. Primikirios, V. Risovannaya, K. A. Roubelakis-Angelakis, H. Snoussi, P. Sotiri, S. Tamhankar, P. This, L. Troshin, J. M. Malpica, F. Lefort, J.M. Martinez-Zapater, *Mol Ecol*., **15(12)**:3707-14, (2006) DOI: 10.1111/j.1365-294X.2006.03049.x
5. T. Beridze, I. Pipia, J. Beck, S.-C. Hsu, M. Gamkrelidze, M. Gogniashvili, V. Tabidze, P. This, R. Bacilieri, V. Gotsiridze, M. Glonti, B. School, *Bull. Georg. Natl. Acad. Sci*., vol. **5**, no 1, pg. 98-103, (2011)
6. V. Tabidze, G. Baramidze, I. Pipia, M. Gogniashvili, L. Ujmajuridze, T. Beridze, A. G. Hernandez and B. Schaal, (2014). *OENO One*, **48**, (2014)
7. G. Zecca, F. Grassi, V. Tabidze, I. Pipia, N. Kotaria, A. Kotorashvili, T. Beridze, *Curr. Genet.*, **66**, (2019) DOI 10.1007/s00294-019-01004-7
8. I. Pipia, M. Gogniashvili, V. Tabidze, T. Beridze, M. Gamkrelidze, V. Gotsiridze, G. Melyan, M. Musayev, V. Salimov, J. Beck, B. Schaal, *Vitis*, **51**(3), (2012)
9. I. Pipia, N. Kunelauri, M. Gogniashvili, N. Kotaria, A. Kotorashvili, T. Lacombe, V. Tabidze, *Acta Horticult*., (2017), DOI:10.17660/ActaHortic.2017.1188
10. I. Pipia, L. Roma and V. Tabidze, *Acta Hortic*. 1248, **235-240**, (2019), DOI: 10.17660/ActaHortic. 2019.1248.34
11. I. Pipia, M. Bezhashvili, A. Gujabidze, T. Nozadze, N. Zhorzholiani, V. Tabidze, *Bull. Georg. Natl. Acad. Sci.*, Vol. **17**, No 2, (2023)
12. I. Pipia, T. Lacombe, A. Kotorashvili, N. Kotaria & V. Tabidze, *Vitis*, **62(4),** (2023) https://doi.org/10.5073/vitis.2023.62.193-201
13. I. Pipia, L. Ujmajuridze, L. Mamasakhlisashvili, A. Hernandez and V. Tabidze, *OENO One*, *in press* (2024)

14. X. Huang, W. Tan, F. Li, R Liao, Z. Guo, T. Shi, Z. Gao, *Gene*, Vol. **793**, (2021) https://doi.org/10.1016/j.gene.2021.145751.

15. A. Lodhi, G. Ye, N. F. Weeden, B. I. Reisch, *Plant Mol. Biol. Repor.*, **12:6-13**, (1995) DOI: 10.1007/BF02668658

16. A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, G.J., *Bioinform.*, **25 (9)** 1189-1191, (2009) DOI: 10.1093/bioinformatics/btp033