საქართველოს მეცნიერებათა ეროვნული აკადემიის მოამბე, ტ. 17, №2, 2023
BULLETIN OF THE GEORGIAN NATIONAL ACADEMY OF SCIENCES, vol. 17, no. 2, 2023

Molecular Genetics

Complete Plastid Genomics of Wild Grapevines (*Vitis vinifera* ssp. *sylvestris*): Next Generation Sequencing, Comparative Genomics, Phylogenetics and Genome Annotation

Ia Pipia*, Mari Bezhashvili*, Ana Gujabidze*, Tengiz Nozadze*, Natia Zhorzholiani*, Vazha Tabidze*

*Institute of Molecular Genetics, Agricultural University of Georgia, Tbilisi, Georgia

(Presented by Academy Member Tengiz Beridze)

The complete chloroplast DNA sequence of wild grapevine (*Vitis vinifera* ssp. *sylvestris*) samples of Europe (Italy, Spain, Portugal, France, Germany) and Mediterranean basin (Morocco, Turkey) were analyzed by Illumina sequencing. Next-generation sequencing data were subjected to comparative genomics, phylogenetic, and genome annotation studies. It was shown that all analyzed plastoms, except for the one from Turkey, belong to the GTA or Chkhaveri-Pinot noir haplotype, while the Turkish sample belongs to the ATA or Chardonnay-Meskhuri mtsvane haplotype. Genome assembly and annotation revealed that the each plastome encode 131 genes including protein coding, tRNA and rRNA genes. Comparative and phylogenetic analyses of plastid DNA sequences suggest that the genetic relationships among the above mentioned wild grapevine samples are consistent with their haplotype affiliations. © *2023 Bull. Georg. Natl. Acad. Sci.*

Vitis vinifera ssp. *sylvestris*, grapevine, plastid DNA, next-generation sequencing, comparative genomics, genome annotation

Wild grapevines (*V. vinifera* ssp. sylvestris) are the progenitors of cultivated grapevines. They first appeared around 65 million years ago, they are predominantly forest climbers, occurring in disjunct populations from the Atlantic coast to Tadzhikistan and the western Himalayas [1]. According to literature, the cultivation of wild grapevines began approximately 6.000-8.000 years ago in the Transcaucasia, between the Black Sea and Iran (Zagros mountains) [2-6]. There are some archeological and palaeobotanical findings pointed Georgia (South Caucasus) as a center of grapevine domestication [1, 2, 7-12]. For the understanding the molecular bases of domestication process and its geographic origin, it is necessary to evaluate of genetic profiles and conduct comparative genomic and phylogenetic studies of both wild and cultivated grapevines.

In this manuscript, we present the results obtained from the Next-Generation Sequencing (NGS), Comparative Genomics, Genome Annotation and Phylogenetic analyses of uniparentally inherited plastid genomes of wild grapevines from Europe (Italy, Spain, Portugal, France, Germany) and Mediterranean basin (Morocco, Turkey). Unlike nuclear genomes, plastid genomes have low rate of point mutations and conserved molecular structures, making them extensively used in the studies of phylogenetic lineages.

Materials and Methods

Cuttings of wild grapevines from Europe and Mediterranean basin were received from the Grapevine collection of Vassal-Montpellier at the National Research Institute for Agriculture, Food and Environment (INRAE), France. The cuttings were grown to obtained young green leaves for DNA extraction. Leaves were ground in liquid nitrogen to homogenize them, and DNA extraction was conducted using CTAB-based protocol [13]. The NGS-reactions were carried out in the Roy J. Carver Biotechnology Center, University of Illinois at Urbana Champaing (UIUC). The construction of shotgun genomic DNA libraries and sequencing on an Illumina HiSeq have been described in our previous publication [14].

Results and Discussion

In the frame of the presented research wild grapevines of Italy, Spain, Portugal, France, Germany Morocco and Turkey (Table 1) were subject to NGS, Comparative Genomics study, Genome Annotation and Phylogenetic Analyses. For all analyzed grapevine samples, the whole plastid genomes were assembled, the genomic regions for the detection of haplotypes were defined and more than 130 genes for each plastome were annotated.

Haplotype affiliations. The NGS results for the whole plastid genome have shown that all analyzed wild grapevine plastomes, except wild grapevine of

Table 1. List of analyzed wild grapevines (*V. vinifera* ssp. *sylvestris*) received from INRA Vassal-Montpellier collection, France

Accession name and code	Country Provenance	Sex	Identification Status
V. silvestris Orroa desa ide 3 (8500Mtp436)	Italy	Male	Natural, wild
Vitis sylvestris J 2-3 Jaen (8500Mtp378)	Spain	Male	Natural, wild
V. silvestris Portugal 110104 (3) (8500Mtp421)	Portugal	Female	Natural, wild
Lambrusque Gresigne 2 (8500Mtp125)	France	Male	Natural, wild
Vitis silvestris Ketsch 27 (8500Mtp34)	Germany	Female	Natural, wild
Lambrusque Akchour S21 Ind 4 (8500Mtp241)	Morocco	Male	Natural, wild
Lambrusque de semis Cirali 1-20 (23433Mtp1-20)	Turkey	-	Wild

Table 2. Plastid genome lengths, haplotypes and Comparative analysis of analyzed wild grapevines. The names of countries indicate the geographical origin of the samples

Samples	Length, bp	Haplotype	Comparative analysis with reference DNA
Italy	160928	GTA	100% homology with Maxxa
Spain	160928	GTA	100% homology with Maxxa
Portugal	160928	GTA	100% homology with Maxxa
France	160924	GTA	4 Gaps difference with Maxxa
Germany	160928	GTA	100% homology with Maxxa
Morocco	160926	GTA	3 SNPs and 6 gaps difference with Maxxa
Turkey	160905	ATA	14 SNPs and 14 gaps difference with Meskhuri mtsvane

Turkey, belong to the GTA or Chkhaveri-Pinot noir plastid haplotype. The wild grapevine of Turkey belongs to the ATA or Chardonnay-Meskhuri mtsvane plastid haplotype. Table 2 presents the plastid genome lengths, haplotype affiliations, and comparative genome analysis for the each studied wild grapevines. It should be underlined that mentioned haplotypes (GTA and ATA) were derived from our earlier studies were set of Worldwide cultivated grapevines and South Caucasian wild grapevine samples were analyzed by point substitutions found in the three polymorphic regions (intergenic spacers trnH-psbA and aCCD-psaI, intron rpl16) of chloroplast genomes [14,15]. The first letter in the haplotype acronyms correspond to the nucleotide substitution in the position of 205 bp of trnH-psbA intergenic spaces, and the second and third letters correspond to the substitutions found in two positions of rpl16 intron at 86715 bp and 86721 bp. Genome positions of nucleotide substitutions are defined according to reference Maxxa (*Vitis vinifera* ssp. *vinifera*) [16].

Interestingly, that obtained results are completely coincides with our previous observations, which demonstrated the dominance of GTA haplotype in wild grapevine of Europe and Mediterranean basin [17], as well as the presence of ATA haplotype in the wild grapevine of Turkey [14]. It should be underlined that our results also confirm the presence of all four haplotypes (AAA – Rkatsiteli Haplotype, ATT – Saperavi-Cabernet Sauvignon haplotype, ATA – Chardonnay-Meskhuri mtsvane and GTA – Chkhaveri-Pinot noir haplotypes) only in South Caucasia which indicates that this geographic region is likely the center of grapevine domestication.

Genome assembly and annotation. For the assembly of grapevine plastid genomes SOAPdenovo computer program was used. As a reference genome for wild grapevines of GTA haplotype Maxxa genome (GenBank DQ424856.1) was used, which has the same GTA haplotype. For the wild grapevine from Turkey, plastid genome of Georgian cultivated grapevine variety Meskhuri mtsvane was used as a reference as both samples belong to ATA haplotype. After the genome assembly by which the exact length of the each plastid genome was determined (see Table 2) gene annotations were conducted. Sequencing analyses indicated that plastid genome encodes 131 genes including 85 protein coding genes, 38 tRNA genes and 8 rRNA genes.

The part of wild grapevine chloroplast genomes from this study have been deposited in the DNA Data Bank of Japan. Accession number for grape sample from Italy is LC715366.1, for Spain LC721283.1, and for Portugal LC715421.1.

Comparative Genomics and Phylogenetics. For the demonstration of genetic relationships and linkage among studied wild grapevines phylogenetic tree of corresponding complete chloroplast genomes were subject to Average Distance Analysis by MAFT jalview program (Fig. 1).

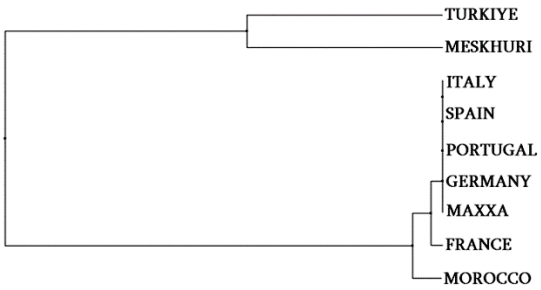

Fig. 1. Phylogenetic tree based on Average Distance Analysis of wild grape plastid DNA by MAFT jalview program (https://www.jalview.org/).

The phylogenetic tree contains all studied plastomes and their reference genomes. Two independent clades are distinguished on the tree. One of them is presented only by the members of GTA haplotype – the wild grapevine samples of Italy, Spain, Portugal, France, Germany Morocco and cultivar Maxxa. The second clade is presented only by samples of ATA haplotype – the wild grapevine from Turkey and reference Georgian

cultivar Meskhuri mtsvane. We observed a well-supported grouping among the members of GTA haplotype samples, where most of them reflect 100% identity with the Maxxa reference DNA. Comparative analysis of plastid DNA sequence revealed small difference only for the samples from Morocco and France, where several SNPs and gaps were observed (Table 2).

This research was supported by Shota Rustaveli National Science Foundation of Georgia (SRNSFG) [grant number: FR-19-14378].

მოლეკულური გენეტიკა

ველური ვაზის სრული პლასტიდური გენომიკა (*Vitis vinifera* ssp. *sylvestris*): ახალი თაობის სეკვენირება, შედარებითი გენომიკა, ფილოგენეტიკა და გენომის ანოტაცია

ი. ფიფია*, მ. ბეჟაშვილი*, ა. გუჯაბიძე*, თ. ნოზაძე*, ნ. ჟორჟოლიანი*, ვ. ტაბიძე*

* *საქართველოს აგრარული უნივერსიტეტი, მოლეკულური გენეტიკის ინსტიტუტი, თბილისი, საქართველო*

(წარმოდგენილია აკადემიის წევრის თ. ბერიძის მიერ)

ევროპისა (იტალია, ესპანეთი, პორტუგალია, საფრანგეთი, გერმანია) და ხმელთაშუაზღვის-პირეთის (მაროკო, თურქეთი) ველური ვაზის (*Vitis vinifera* ქვესახ. *sylvestris*) სრული ქლოროპლასტური დნმ-თანმიმდევრობები შესწავლილ იქნა ილუმინას სეკვენირებით. ახალი თაობის სეკვენირების შედეგები გაანალიზდა შედარებითი გენომიკის მიდგომებით. ასევე, შესწავლილ იქნა ველური ვაზის პლასტომების ფილოგენეტიკური კავშირები და ანოტირებულ იქნა მათი გენომები. ნაჩვენებია, რომ შესწავლილი პლასტომების უმრავლესობა მიეკუთვნება GTA, ანუ, ჩხავერი-პინო ნუარის ჰაპლოტიპს. მხოლოდ თურქეთის ველური ვაზის ნიმუში წარმოადგენდა განსხვავებულ, კერძოდ, ATA, ანუ, შარდონე-მესხური მწვანის ჰაპლოტიპს. გენომის ასამბლირებამ და ანოტაციამ თითოეულ საკვლევ პლასტომში 131 გენის (მათ შორის ცილა მაკოდირებელი, სატრანსპორტო რნმ-ისა და რიბოსომული გენების) არსებობა აჩვენა. შესწავლილი გენომების შედარებითი გენომიკის მიდგომებითა და ფილოგენეტიკური ანალიზით ნაჩვენებია, რომ შესწავლილ ნიმუშებს შორის არსებული გენეტიკური კავშირები მათი ჰაპლოტიპების შესაბამისია.

REFERENCES

1. Zohary D., Horf M. (2000) Domestication of plants in the Old World. Oxford University Press, Oxford UK.
2. McGovern P. E. (2003) Ancient wine: the search for the origins of viniculture. Princeton University Press, Princeton.
3. Barnard H., Dooley A. N., Areshian G. et al. (2011) Chemical evidence for wine production around 4000 BCE in the late chalcolithic near eastern highlands, *J. Archaeol. Sci.,* 38: 977-984.
4. Myles S., Boyko A.R., Owens C.L. et al. (2011) Genetic structure and domestication history of the grape, *Proc. Natl. Acad. Sci.,* USA **108**(9): 3530-5.
5. Jackson R.S. (1994) Wine science principles and application, pp 2-4. Academic Press, New York.
6. Grassi F., Arroyo-Garcia R. (2020) Origins and domestication of the grape, *Front. Plant Sci.,* 11.
7. Laucou V., Lacombe T., Dechesne F. et al. (2011) High throughput analysis of grape genetic diversity as a tool for germplasm collection management, *Theoretical and Applied Genetics,* 122: 1233-1245.
8. Imazio S., Maghradze D., De Lorenzis G. et al. (2013) From the cradle of grapevine domestication: molecular overview and description of Georgian grapevine (*Vitis vinifera* L.) germplasm, *Tree Genetics & Genomes,* 9: 641-658.
9. Ekhvaia J., Gurushidze M., Blattner F.R. (2014) Genetic diversity of Vitis vinifera in Georgia: relationships between local cultivars and wild grapevine, *V. vinifera* subsp. *Sylvestris. Genet. Resour. Crop. Evol.*, 61: 1507-1521.
10. De Lorenzis G., Chipashvili R., Failla R.O., Maghradze D. et al. (2015) Study of genetic variability in *Vitis vinifera* L. germplasm by high-throughput Vitis18kSNP array: the case of Georgian genetic resources, *BMC Plant Biol.,* 15: 154.
11. Laucou V., Laulay A., Bacilieri R. et al. (2018) Extended diversity analysis of cultivated grapevine Vitis vinifera with 10K genome-wide SNPs, *PLOS One,* 13: 0192540.
12. Sargolzaei M., Rustioni L., Cola G. et al. (2021) Georgian grapevine cultivars: ancient biodiversity of future viticulture, *Frontiers in Plant Science*, 12: 630122.
13. Lodhi M. A., Ye G.N, Weeden N. F., Reisch B. I. (1994) A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species, *Plant Molecular Biology Reporter,* 12: 6-13.
14. Pipia I., Gogniashvili M., Tabidze V. et al. (2012) Plastid DNA sequence diversity in a worldwide set of grapevine cultivar (*Vitis vinifera* subsp. *sylvestris*), *Vitis* 51: 119-124.
15. Beridze T., Pipia I., Beck J. et al. (2011) Plastid DNA sequence diversity in a worldwide set of grapevine cultivars (*Vitis vinifera* L. subsp. *vinifera*), *Bull. Georg. Natl. Acad. Sci.,* **5**(1) 98-103.
16. Jansen R. K., Kaittanis C., Saski C. et al. (2006) Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids, *BMC Evolutionary Biology*, 6:32.
17. Pipia I., Kunelauri N., Gogniashvili M. et al. (2017) Complete plastid genomes of South Caucasian, European and Mediterranean basin wild grapevines (*Vitis vinifera* subsp. *sylvestris*), *Acta Horticulturae,* 1188: 351-356.

Received March, 2023