Turczaninowia 27, 1: 124–129 (2024)
DOI: 10.14258/turczaninowia.27.1.16
http://turczaninowia.asu.ru

ISSN 1560–7259 (print edition)
TURCZANINOWIA
ISSN 1560–7267 (online edition)

УДК 582.783.2+575.113(282.247.36)

The complete chloroplast genome sequence of *Vitis vinifera* 'Krasnostop Zolotovskiy', an autochthonous variety of the Don Valley

F. S. Sharko[1, 3]*, K. O. Petrova[1, 4], N. V. Molchanov[2, 5], M. V. Patrushev[1, 6], D. Y. Fedosov[1, 7], S. V. Toshchakov[1, 8]

[1] *National Research Center "Kurchatov Institute", Ploshchad Akademika Kurchatova, 1, Moscow, 123182, Russian Federation*

[2] *All-Russian non-governmental organization Russian Geographical Society, Novaya Ploshchad, 10, bldg 2, Moscow, 109012, Russian Federation*

[3] *E-mail: fedosic@gmail.com; ORCID iD: https://orcid.org/0000-0002-1189-5597*

[4] *E-mail: petrova_ko@nrcki.ru; ORCID iD: https://orcid.org/0000-0003-1909-2171*

[5] *E-mail: dimitri.kovalev@gmail.com; ORCID iD: https://orcid.org/0009-0002-2066-7665*

[6] *E-mail: me@maxpatrushev.ru; ORCID iD: https://orcid.org/0000-0002-4748-9283*

[7] *E-mail: fedosov_dy@nrcki.ru; ORCID iD: https://orcid.org/0000-0002-6833-9233*

[8] *E-mail: stepan.toshchakov@gmail.com; ORCID iD: https://orcid.org/0000-0001-7549-3450*

* *Corresponding author*

Keywords: assembly, de novo, genome annotation, grapevine varieties, phylogenetic analysis.

Summary. DNA sequence data from complete chloroplast genomes expand our understanding of plant diversity and contributes to our understanding of evolutionary relationships among species. *De novo* assembly and annotation of plant chloroplast genomes are difficult due to the high number of repeats and low number of plant chloroplast genomes in international databases. The development of long-read sequencing platforms such as PacBio and Oxford Nanopore, as well as specialized assemblers, is simplifying the process of assembling and annotating large circular molecules. In this work, we report the complete chloroplast genome of the autochthonous *Vitis vinifera* variety 'Krasnostop Zolotovskiy', obtained *de novo* by the hybrid assembly method using short and long DNA reads. The circular genome of the chloroplast is 160,927 bp long with a GC ratio of 37.38 %. It has four subregions: large single-copy (LSC) region of 71,946 bp, a small single-copy (SSC) region of 36,273 bp, and two inverted repeat regions (IRs) of 26,354 bp each. The chloroplast genome of *V. vinifera* 'Krasnostop Zolotovskiy' harbours 129 genes, comprising 84 protein-coding genes, 8 rRNA genes, and 37 tRNA genes. We also carry out a phylogenetic analysis based on all the complete sequences of grape chloroplast genomes available in the databases, which shows that the *V. vinifera* 'Krasnostop Zolotovskiy' does not have a direct origin with any European variety.

Полная последовательность хлоропластного генома *Vitis vinifera* 'Красностоп Золотовский', автохтонного сорта из долины Дона

Ф. С. Шарко[1], К. О. Петрова[1], Н. В. Молчанов[2], М. В. Патрушев[1], Д. Ю. Федосов[1], С. В. Тощаков[1]

1 Национальный исследовательский центр «Курчатовский институт», пл. Академика Курчатова, д. 1, г. Москва, 123182, Россия

2 Общероссийская общественная организация «Русское географическое общество», Новая площадь, д. 10, стр. 2, г. Москва, 109012, Россия

Ключевые слова: аннотация генома, де ново, сборка, сорта винограда, филогенетический анализ.

Поступило в редакцию 27.10.2023
Принято к публикации 19.04.2024

Submitted 27.10.2023
Accepted 19.04.2024

Аннотация. Данные о ДНК-последовательностях полных хлоропластных геномов расширяют наше представление о разнообразии растений и способствуют пониманию эволюционных взаимоотношений между видами. Сборка *de novo* и аннотация геномов хлоропластов растений затруднены из-за большого количества повторов и небольшого количества хлоропластных последовательностей ДНК растений в международных базах данных. Разработка платформ для секвенирования с длинными чтениями, таких как PacBio и Oxford Nanopore, а также специализированных ассемблеров упрощает процесс сборки и аннотирования больших кольцевых молекул. В данной работе приводится полный геном хлоропласта автохтонного сорта винограда *Vitis vinifera* 'Красностоп Золотовский' из долины Дона, полученный *de novo* методом гибридной сборки с использованием коротких и длинных чтений ДНК. Кольцевой геном хлоропласта имеет длину 160927 пар оснований с соотношением GC 37,38 %. Геном состоит из четырех субрегионов: большой области с одной копией (LSC) длиной 71946 п. н., небольшой области с одной копией (SSC) длиной 36273 п. н. и двух областей инвертированных повторов (IR) длиной 26354 п. н. каждая. Хлоропластный геном сорта *V. vinifera* 'Красностоп Золотовский' содержит 129 генов, в том числе 84 гена, кодирующих белки, 8 генов рРНК и 37 генов тРНК. Также проведён филогенетический анализ на основе всех имеющихся в базах данных полных последовательностей хлоропластных геномов винограда, который показал, что *V. vinifera* 'Красностоп Золотовский' не имеет прямого происхождения ни от одного европейского сорта.

Introduction

The evolutionary history of autochthonous varieties of *Vitis vinifera* L. is of great interest due to their unique adaptation features to local climate peculiarities. Through long-term natural selection, these varieties have developed genetic determinants of advantageous traits that enable them to adjust to the climatic conditions and circadian rhythms specific to the area (Labagnara et al., 2018). The focus of numerous viticultural researchers currently centers on the potential of autochthonous varieties. They aim to investigate and utilize their benefits in light of shifting environmental conditions as well as a constantly changing array of plant pests. 'Krasnostop Zolotovskiy' (Fig. 1) is a crucial autochthonous variety for contemporary Russian winemaking and is revered by several experts as the finest local grape variety within the country (Robinson et al., 2013). The variety comes from the Don River valley (actually Rostov Oblast in Southern Russia) and was first mentioned by its name in 1814. It was previously assumed that all Don Valley varieties were imported from Western Europe and the Balkans. However, recent studies based on single nucleotide polymorphism (SNP) analysis show no direct relationship with Caucasian and Balkan varieties (Fedosov et al., 2021). Phylogenetic analysis of the chloroplast genome of *V. vinifera* 'Krasnostop Zolotovskiy' can offer significant insights into the evolutionary processes of grapevine varieties (Ballas, 1877).

Fig. 1. Plant image of *Vitis vinifera* 'Krasnostop Zolotovskiy', Sikory Estate, Novorossiysk, Krasnodar Region, Russia, August 2023 (photo by N. Molchanov).

Materials and methods

Fresh leaves of *V. vinifera* 'Krasnostop Zolotovskiy' were collected in the vineyards of Southern Russia from Sikory Estate, Novorossiysk, Krasnodar Region (44°53'9.9"N, 37°37'8.1"E) by Nikolay Molchanov. A specimen was deposited at the National Research Center "Kurchatov Institute" (Dmitriy Fedosov, dimitri.kovalev@gmail.com) under the voucher number LIBV0013. The DNA extraction protocol used in this research was adjusted from that of (Sandra Lo Piccolo, 2012). Modifications included performing all centrifugation procedures at 3500 g and 4 °C. The centrifugation time was 15 min and 45 min for DNA extraction with chloroform-isoamyl alcohol and isopropanol precipitation of DNA, respectively. Prior to sequencing on a GridION device (Oxford Nanopore Technologies, UK), an extra purification step using Genomic Tip 20/G columns (Qiagen, Germany) was performed according to the manufacturer's instructions (Sharko et al.,

2021). The concentration and quality of extracted DNA was assessed using a Nanodrop 1000 device (Thermo Fischer Scientific, Waltham, MA, USA) and a Qubit fluorometer with the Qubit™ dsDNA BR Assay Kit (Thermo Fischer Scientific, Waltham, MA, USA). DNA fragment libraries were prepared for Illumina sequencing using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England BioLabs, Ipswich, MA, USA) according to the manufacturer's protocol. The resulting libraries were sequenced on the NovaSeq 6000 platform (Illumina, San Diego, CA, USA) using 2*150 bp paired-end chemistry. Long-read sequencing was performed on a GridION instrument using the Ligation Sequencing Kit according to the manufacturer's guidelines.

Read filtering, trimming of sequencing adapters and low-quality read regions were performed with the fastp tool (Chen et al., 2018). Hybrid *de novo* assembly was performed using the SPAdes v3.15.0 assembler (Bankevich et al., 2012) with a coverage of 7025.5× (Figure S1). The obtained chloroplast genome was annotated using the OGDRAW platform (Greiner et al., 2019). The complete chloroplast genome of *Vitis vinifera* 'Krasnostop Zolotovskiy' was submitted to GenBank (Accession number: OR500062).

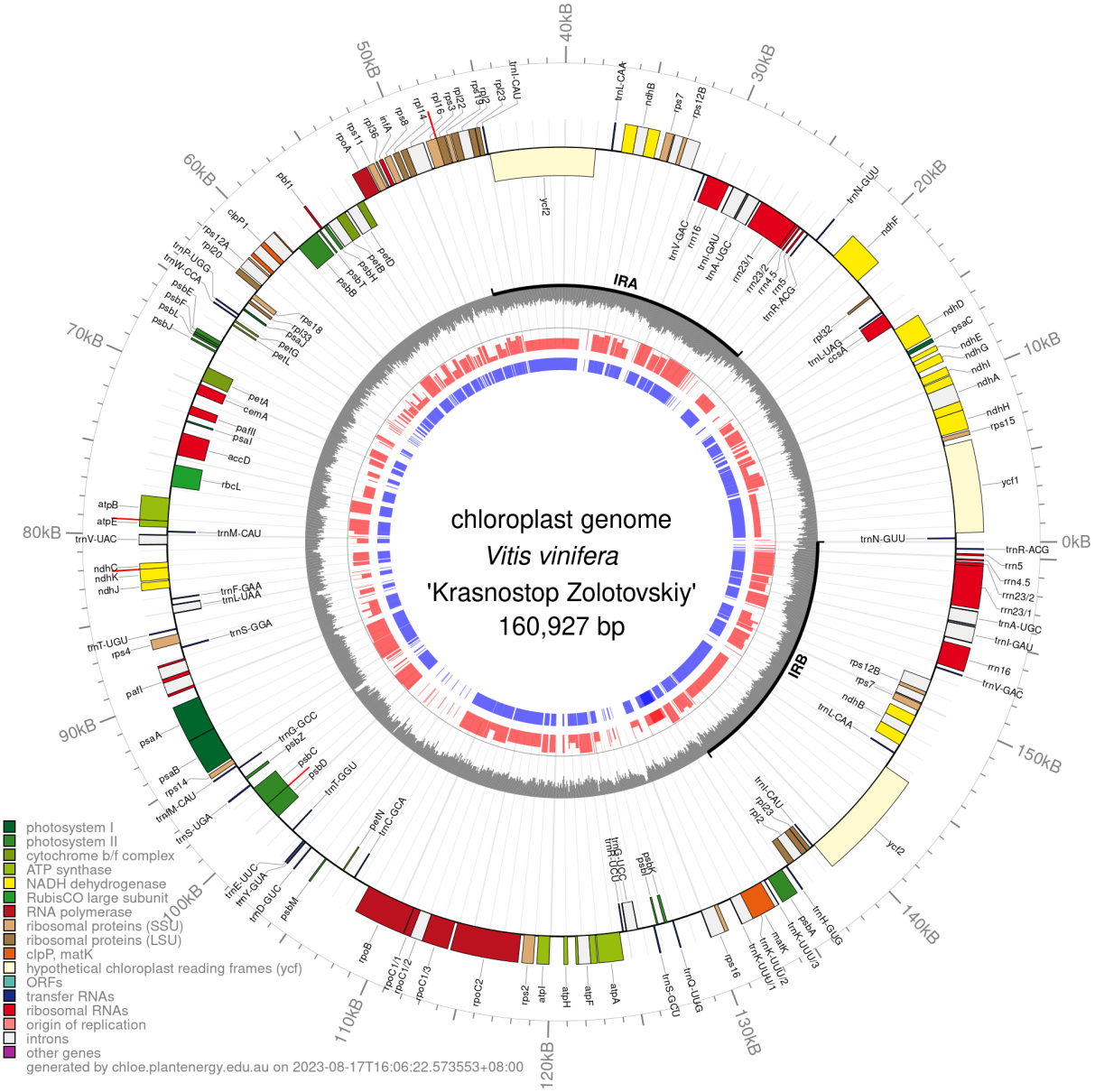

Fig. 2. Complete chloroplast genome map of *Vitis vinifera* 'Krasnostop Zolotovskiy'. Color of coding genes is based on functional groups they belong to. Dark gray color of inner circle indicates GC content.

For the phylogenetic analysis, common conserved genes extracted from 17 *Vitis* chloroplast genomes were aligned using mafft v7.245 (Katoh et al., 2019) with an iterative method (G-INS-i) using the default parameter settings. To construct the maximum likelihood (ML) phylogenetic tree, the RAxML algorithm (Stamatakis, 2015) was employed, and bootstrap values were obtained from 1000 replicates. The resulting tree was visualized using the iTOL service (Letunic, Bork, 2019).

Result

The chloroplast genome of *V. vinifera* 'Krasnostop Zolotovskiy' is 160,927 bp in length (GC ratio: 37.38 %). The genome contains a large single-copy (LSC) region of 71,946 bp, a small single-copy (SSC) region of 36,273 bp, and a pair of inverted repeat regions (IRA and IRB) of 26,354 bp. It possesses 129 genes, including 84 protein-coding genes, 37 tRNA genes, and 8 rRNA genes (Fig. 2). 16 genes were found to be duplicated, including 5 protein-coding genes (*ndhB*, *rpl2*, *rpl23*, *rps7* and *ycf2*), 7 tRNA genes (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG* and *trnV-GAC*) and 4 rRNA genes (*rrn16, rrn23, rrn4.5, rrn5*). 13 protein-coding genes and 8 tRNA genes contain one intron, while *clpP1*, *pafI* and *rps12B* genes contain two introns.

To determine the phylogenetic position of the *V. vinifera* 'Krasnostop Zolotovskiy', a ML tree was constructed using the protein-coding genes of the *Vitis* chloroplast genomes collected in the current study, as well as other 16 members of the *Vitis* family, loaded from the NCBI database, with *V. rotundifolia* as the outgroup. Phylogenetic analysis showed that *Vitis vinifera* 'Krasnostop Zolotovskiy' belongs to the group of Asian varieties *V. vinifera* 'Fengzao' and *Vitis vinifera* 'Kyoho' (Fig. 3) with strong support (BP = 100), indicating a close relationship of *V. vinifera* 'Krasnostop Zolotovskiy' with these types. At the same time, European and Caucasian varieties form two separate groups with equally strong support (BP = 100).

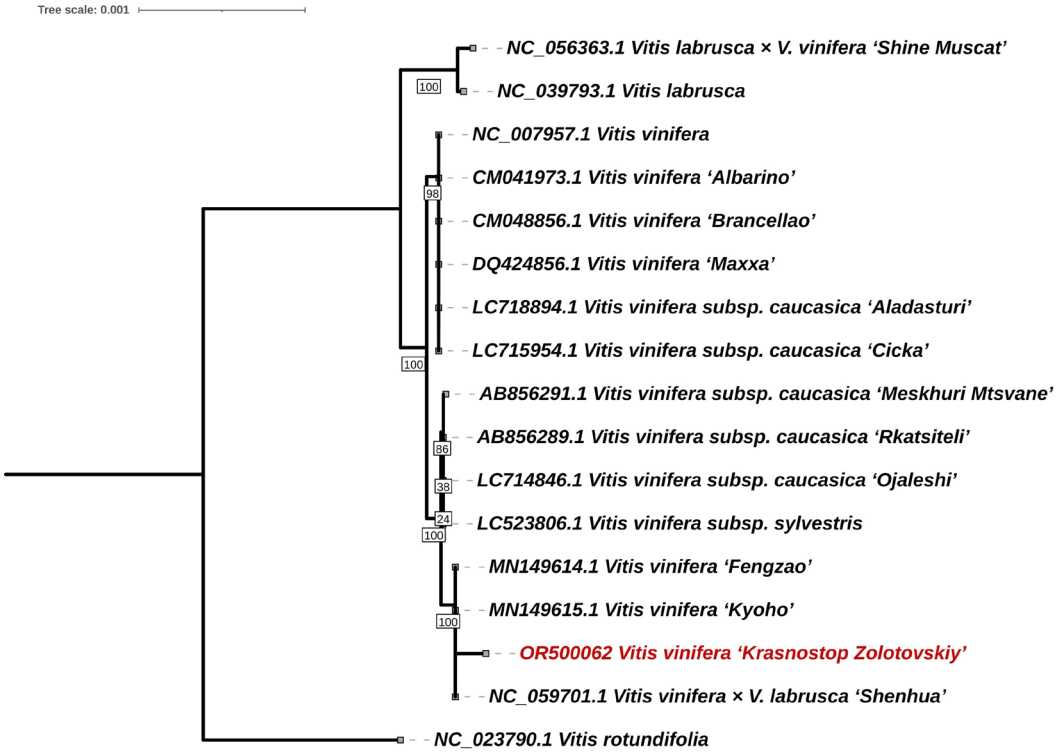

Fig. 3. Phylogenetic position of *Vitis vinifera* 'Krasnostop Zolotovskiy' in the ML phylogenetic tree constructed using sequences of protein-coding genes from 17 complete chloroplast genomes of different grapevines. The following sequences were used in the analysis: *V. vinifera* subsp. *caucasica* 'Meskhuri Mtsvane' (AB856291.1), *V. vinifera* subsp. *caucasica* 'Rkatsiteli' (AB856289.1), *V. vinifera* subsp. *caucasica* 'Ojaleshi' (LC714846.1), *V. vinifera* subsp. *sylvestris* (LC523806.1), *V. vinifera* 'Maxxa' (DQ424856.1), *V. vinifera* subsp. *caucasica* 'Cicka' (LC715954.1), *V. vinifera* (NC_007957.1), *V. vinifera* subsp. *caucasica* 'Aladasturi' (LC718894.1), *V. vinifera* 'Fengzao' (MN149614.1), *V. vinifera* 'Krasnostop Zolotovskiy' (OR500062), *V. vinifera* 'Kyoho' (MN149615.1) , *V. vinifera* 'Brancellao' (CM048856.1) (unpublished), *V. vinifera* 'Albarino' (CM041973.1) (unpublished), *V. rotundifolia* (NC_023790.1), *V. labrusca* × *V. vinifera* 'Shine Muscat' (NC_056363.1), *V. labrusca* (NC_039793.1) and *V. vinifera* × *V. labrusca* 'Shenhua' (NC_059701.1) (Jansen et al., 2006; Wen et al., 2018; Pipia et al., 2019; Guo et al., 2020a, b; Wu et al., 2020; Zhang et al., 2021). Bootstrap support values are shown on the nodes.

Discussion and conclusion

This study presents the complete assembly and annotation of the chloroplast genome sequence of the *V. vinifera* 'Krasnostop Zolotovskiy'. Based on phylogenetic analysis (Fig. 3) using protein-coding genes in the publicly available complete chloroplasts genomes, it was found that 'Krasnostop Zolotovskiy' has no direct genetic relationship with other Caucasian and European varieties of *V. vinifera*. Instead, it forms a distinct cluster in the genetic dendrogram of *V. vinifera* along with other autochthonous varieties, which strongly contradicting earlier theories suggesting its origin from Eastern European varieties introduced into Russia. We cannot unambiguously confirm the results obtained in the work of Fedosov et al. (2021), as Asian varieties were not previously represented in that study. But it can be assumed that analysis of chloroplast genomes can become an alternative tool for studying the genetic relationships of grape varieties. However, to test the applicability of this method, it is necessary to use extended samples and analyze the correlations of chloroplast analyzes with correlations of SNPs. The exploration of the chloroplast genome of the *V. vinifera* 'Krasnostop Zolotovskiy' contributes to our understanding the evolution and genetic diversity of grapevine cultivars. Further investigation of the genetic composition and structure of the chloroplast genome may provide additional insights into the origins of unique traits and the domestication process (Arroyo-García et al., 2006). In addition, the information gained from studying the chloroplast genome could be used to improve agricultural practices and obtain new vine varieties with improved characteristics.

Data availability statement

Chloroplast genome sequence data that support the findings of this study are openly available in GenBank of NCBI (https://www.ncbi.nlm.nih.gov/nuccore/OR500062) under the accession no. OR500062. The associated BioProject, SRA, and Bio-Sample numbers were PRJNA1011053, SRR25823228 and SAMN37209290, respectively.

Acknowledgements

The work was carried out within the state assignment of NRC "Kurchatov institute".

REFERENCES / ЛИТЕРАТУРА

Arroyo-García R., Ruiz-García L., Bolling L., Ocete R., López M. A., Arnold C., et al. 2006. Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphisms. *Molec. Ecol.* 15(12): 3707–3714. DOI: 10.1111/j.1365-294X.2006.03049.x

Ballas M. 1877. Essay on Winemaking in Russia (The Caucasus and Crimea). *Historical and Statistics Collection on Winemaking* 1: 137–138.

Bankevich A., Nurk S., Antipov D., Gurevich A., Dvorkin M., Kulikov A., et al. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Computat. Biol.* 19(5): 455-477. DOI: 10.1089/cmb.2012.0021

Chen S., Zhou Y., Chen Y., Gu J. 2018. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34(17): i884–i890. DOI: 10.1093/bioinformatics/bty560

Fedosov D. Y., Korzhenkov A. A., Petrova K. O., Sapsay A. O., Sharko F. S., Toshchakov S. V., Kolosova A. A., Bakhmutova E. D., Patrushev M. V. 2021. SNP-based analysis reveals authenticity and genetic similarity of Russian indigenous *Vitis vinifera* grape cultivars. *Plants* 10(12): 1–12. DOI: 10.3390/plants10122696

Greiner S., Lehwark P., Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) Version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucl. Acids Res.* 47(W1): W59–W64. DOI: 10.1093/nar/gkz238

Guo L., Guo S., Wang C., Guo D. 2020a. The complete chloroplast genome sequence of *Vitis vinifera* 'Feng Zao'. *Mitochondrial DNA Part B* 5(2):1789–1790. DOI: 10.1080/23802359.2020.1749167

Guo L., Wang C., Guo S., Guo D. 2020b. The complete chloroplast genome sequence of *Vitis vinifera* 'Kyoho'. *Mitochondrial DNA Part B* 5(2): 1787–1788. DOI: 10.1080/23802359.2020.1749166

Jansen R., Kaittanis C., Saski C., Lee S., Tomkins J., Alverson A., Daniell H. 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(April): 32. DOI: 10.1186/1471-2148-6-32

Katoh K., Rozewicki J., Yamada K. 2019. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* 20(4): bbx108. DOI: 10.1093/bib/bbx108

Labagnara T., Bergamini C., Caputo A., Cirigliano P. 2018. *Vitis vinifera* L. germplasm diversity: A genetic and ampelometric study in ancient vineyards in the South of Basilicata region (Italy). *VITIS – Journal of Grapevine Research* 57: 1–8 DOI: 10.5073/vitis.2018.57.1-8

Letunic I., Bork P. 2019. Interactive tree of Life (ITOL) v4: Recent updates and new developments. *Nucl. Acids Res.* 47(W1): W256–W259. DOI: 10.1093/nar/gkz239

Pipia I., Roma L., Tabidze V. 2019. Annotation of whole-plastid genomes of wild grapes (*Vitis vinifera* subsp. s*ylvestris*) from the South Caucasus, Europe and the Mediterranean Basin. *Acta Hort.* 1248(August): 235–240. DOI: 10.17660/ActaHortic.2019.1248.34

Robinson J., Harding J., Vouillamoz J. 2013. Wine grapes: a complete guide to 1,368 vine varieties, including their origins and flavours. *Penguin UK.* DOI: 10.1017/jwe.2013.9

Sandra Lo Piccolo 2012. A simple and rapid DNA extraction method from leaves of grapevine suitable for polymerase chain reaction analysis. *African J. Biotechnol.* 11(45): 10305–10309. DOI: 10.5897/AJB11.3023

Sharko F., Gladysheva-Azgari M., Tsygankova S., Mitrofanova I., Boulygina E., Slobodova N., Smykov A., Rastorguev S., Nedoluzhko A. 2021. The complete chloroplast genome sequence of cultivated *Prunus persica* cv. 'Sovetskiy'. *Mitochondrial DNA. Part B, Resources* 6(10): 2882–2883. DOI: 10.1080/23802359.2021.1972861

Stamatakis A. 2015. Using RAxML to infer phylogenies. *Current Protocols in Bioinformatics* 51(September): 6.14.1-6.14.14. DOI: 10.1002/0471250953.bi0614s51

Wen J., Harris A., Kalburgi Y., Zhang N., Xu Y., Zheng W., Ickert-Bond S., Johnson G., Zimmer E. 2018. Chloroplast phylogenomics of the New World grape species (*Vitis*, Vitaceae). *J. Syst. Evol.* 56(4): 297–308. DOI: 10.1111/jse.12447

Wu Y., Zhang W., Song S., Xu W., Zhang C., Ma C., Wang L., Wang S. 2020. Evolution of volatile compounds during the development of Muscat grape 'Shine Muscat' (*Vitis labrusca* × *V. vinifera*). *Food Chemistry* 309(March): 125778. DOI: 10.1016/j.foodchem.2019.125778

Zhang L., Guo D., Liu M., Dou F., Ren Y., Li D., et al. 2021. The complete chloroplast genome sequence of *Vitis vinifera* × *Vitis labrusca* 'Shenhua'. *Mitochondrial DNA. Part B, Resources* 6(1): 166–167. DOI: 10.1080/23802359.2020.1855605