



Article

Chloroplast Genome Variation and Phylogenetic Relationships of Autochthonous Varieties of *Vitis vinifera* from the Don Valley

F. S. Sharko * , K. O. Petrova, M. V. Patrushev, D. Y. Fedosov and S. V. Toshchakov

National Research Center “Kurchatov Institute”, Academician Kurchatov pl., 1, Moscow 123182, Russia; petrova_ko@nrcki.ru (K.O.P.); me@maxpatrushev.ru (M.V.P.); dimfedosov@yandex.ru (D.Y.F.); stepan.toshchakov@gmail.com (S.V.T.)

* Correspondence: fedosic@gmail.com

Abstract: The autochthonous grape varieties of the Don Valley, situated in southern Russia, constitute a distinctive element of regional cultural heritage. These varieties have been adapted over centuries to the region’s specific local climatic and soil conditions. For the most part, these varieties are not imported from other countries. They are closely related to varieties found in Crimea and the North Caucasus. In this study, we obtained the first complete, unfragmented sequences of the chloroplast genomes of eight autochthonous varieties from the Don Valley and one from Crimea. We also performed a comparative analysis of their genomic features. The size of *Vitis vinifera* chloroplast genome sequences varied from 160,925 to 160,991 bp, depending on the cultivar, with a uniform GC ratio of 37.38%. Each genome consists of four subregions: a single copy region (LSC) ranging from 89,158 to 89,336 bp, a small single copy region (SSC) ranging from 19,070 to 19,073 bp, and a pair of inverted repeat regions (IRa and IRb) in the range of 26,292 to 26,353 bp. The chloroplast genomes of the studied *V. vinifera* varieties contained 130 genes, including 85 protein-coding genes, 8 rRNA genes, and 37 tRNA genes. The sequence divergence analysis has enabled the identification of four highly variable regions, which may be utilized as potential markers for phylogenetic analysis. The analysis revealed the presence of 58 to 61 SSRs and multiple long repeated sequences in the chloroplast genomes of these varieties. The phylogenetic analyses of the sequences obtained and complete chloroplast genomes available from public databases indicated that the majority of autochthonous *V. vinifera* varieties do not have a direct origin from any European variety.

Keywords: *Vitis vinifera*; autochthonous varieties; assembly; annotation genome; de novo; chloroplast genome; phylogenetic analysis



Citation: Sharko, F.S.; Petrova, K.O.; Patrushev, M.V.; Fedosov, D.Y.; Toshchakov, S.V. Chloroplast Genome Variation and Phylogenetic Relationships of Autochthonous Varieties of *Vitis vinifera* from the Don Valley. *Int. J. Mol. Sci.* **2024**, *25*, 9928. <https://doi.org/10.3390/ijms25189928>

Academic Editor: Wenpan Dong

Received: 5 August 2024

Revised: 10 September 2024

Accepted: 12 September 2024

Published: 14 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The grape varieties cultivated in the Don Valley region of southern Russia constitute a distinctive gene pool. For the most part, they are not imported from other countries of the world, being close relatives of varieties found in Crimea and the Northern Caucasus [1]. These varieties were historically distributed across numerous villages situated along the right bank of the Don River. At the beginning of the 20th century, over 13,000 hectares of small family vineyards were located between Tsimlyansk and Aksay, spanning a distance of 200 km. Agricultural reforms during the Soviet era have led to the decline of Cossack family winemaking traditions. The varietal composition of the vineyards has also changed: preference was given to hybrid varieties [2], which were obtained by crossing *Vitis vinifera* with *Vitis amurensis* and American species.

Consequently, of the 42 varieties of the Don Valley described in previous publications [3,4], only six are widely represented. The most popular variety is Krasnostop Zolotovskiy. The rest can be found in private and public collections in southern Russia [1]. The genomic studies of indigenous varieties of the Don Valley, which have existed in this

territory for centuries, can serve as a scientific basis to revive the interest of winemakers. The varieties in question are drought- and frost-resistant, and are well-suited to the local climate.

Krasnostop Zolotovskiy is a mid-ripening grape variety from the Black Sea basin [5]. It has small, three-lobed leaves and hermaphroditic flowers. The conical clusters are small and of medium density with dark blue, juicy berries. The ripening period in the Don Valley is 135 days, requiring a Huglin index (HI) value of 2800. It is resistant to fungal diseases and is used for high-quality red wines.

Krasnostop AZOS is an early-to-mid-ripening, interspecific hybrid derived from Krasnostop Zolotovskiy [6]. Its medium-sized, five-lobed leaves are dark green on top and slightly bristly underneath. With 135 days of growth and a 2600 HI, it shows phylloxera tolerance and is known for its juicy, dense-skinned berries.

Kokur Belyi, originating from Crimea [7] and introduced to the Don Valley before 1800, has large, deeply dissected leaves and produces medium-to-large conical clusters. It is a late-ripening variety, taking 160–170 days with 3300–3400 HI used for high-quality white, orange, sparkling, and fortified wines.

Dolgiy Skorospelyi is a separate variety, believed to be a clone of Kokur Belyi, with robust growth and medium-to-large cylindrical clusters. It has a 135–145-day ripening period but is sensitive to frost, fungal diseases, and phylloxera. Grapes are used for fresh consumption and wine production.

Plechistik is an ancestor of several varieties, including Krasnostop Zolotovskiy. It has medium-sized, strongly dissected leaves and functional female flowers. The medium cylindrical clusters produce small-to-medium berries with thin skin, requiring 135–140 days and 2800–2900 HI for ripening. It needs a pollinator like Tsimlyanskiy Chernyi to produce still and sparkling red wines.

Varyushkin is a Black Sea basin grape variety [8] with medium-sized, deeply dissected leaves and hermaphroditic flowers. It has conical clusters of medium density and rounded, black berries with thick skin and juicy flesh. The ripening period is 140–150 days, requiring about 3000 HI. It is frost-resistant and relatively resistant to mildew.

Sibirkovyi likely emerged in the 19th century from a cross between the local Bulanyi variety and the Romanian Pukhlyakovskiy Belyi (Coarne Alba) [9]. It has large, deeply dissected, five-lobed leaves and hermaphroditic flowers. The medium, slightly conical clusters produce oval, greenish-white berries with thin skin and juicy, melting flesh. The ripening period is about 130 days with 2650–2700 °C of active temperatures. This variety is prone to fungal diseases and frost, and is used for light still and sparkling wines.

Puklyakovskiy Belyi, a local name for Coarne Alba imported from the Balkans by a Cossack Pukhlyakov in the 19th century, has large, five-lobed, bubbly leaves. It features functionally female flowers and medium-sized cylindrical clusters with large, oval-ovate, greenish-white berries and thick skin. Its ripening period is about 140–145 days, requiring 3000–3200 HI, and its grapes are used fresh or for winemaking.

Kumshatskiy Belyi is related to Krasnostop Zolotovskiy and Plechistik and originates from the Tsimlyansk sub-region of the Don Valley [10]. It has large, deeply dissected, five-lobed leaves with a stretched middle blade, hermaphroditic flowers, and long, thick clusters. The medium-sized, round, greenish-white berries have fragile skin and juicy flesh. This medium–late-ripening variety takes about 180 days and requires 3300–3700 HI, producing aromatic, well-structured wine.

Chloroplasts are the primary organelles responsible for photosynthesis, the process through which plants transform solar energy into chemical energy [11]. The genomes of chloroplasts encode numerous key proteins that play a vital role in this photosynthetic process, such as enzymes and various accessory components [12]. In angiosperms, chloroplast DNA is usually inherited from the mother, whereas in gymnosperms, it is primarily passed down from the father [13]. This distinction makes chloroplast DNA a crucial focus for exploring heredity and genetic variation within plant populations [14]. Research on chloroplast genomes enables us to track evolutionary changes, to examine hybridization

events between species, and to analyze the relationships among various plant varieties and species [15]. Chloroplast genomes comprise genes that play a role in the synthesis of a variety of metabolites [16], such as pigments, vitamins, and other essential compounds [12]. These substances are crucial for numerous cellular processes and significantly influence the color and flavor of plants.

The first sequenced chloroplast genome was that of *Marchantia polymorpha*, which occurred in 1986 [17]. Subsequently, numerous additional chloroplast genome sequences have been obtained. However, not all of these represented a complete sequence, and thus were only suitable for phylogenetic analysis with known limitations. Notwithstanding, specific regions of chloroplast genomes, including the maturase K (*matK*), ribulose biphosphate carboxylase (*rbcL*), and yeast cadmium factor (*ycfL*) genes, have been extensively employed for DNA barcoding (species identification) [18]. However, these relatively short chloroplast markers were not suitable for distinguishing and analyzing the origin of cultivars. Their close relationship required a higher resolution. More extended sequences, specifically complete chloroplast genomes, may be suitable for full phylogenetic analysis. As sequencing technology continues to advance, methods for chloroplast genome sequencing have become more streamlined. Now, the analysis of an increasing number of chloroplast genome sequences has become more accessible [19,20]. The sequencing of complete chloroplast genomes was initially employed for the identification of cocoa cultivars [21]. Further applications of this approach permitted us to analyze the origin of varieties of significant crops, including hemp, date palm, tomato, rice, and other cultivars [22–25].

In this study, we present a phylogenetic analysis and comparative analysis of genomic features of complete, unfragmented sequences of the chloroplast genomes of eight autochthonous grape varieties from the Don Valley and one from Crimea. These sequences may be used in various fields, including molecular breeding, phylogenetics, and the conservation of endangered species [19,26,27].

2. Results

2.1. Subsection Characterization of the CP Genome Structure of *Vitis vinifera* Variety

The length of the *V. vinifera* chloroplast genome sequences exhibited a range of 66 bp with a variation of 160,925 bp to 160,991 bp among different samples. Each chloroplast genome exhibited the characteristic quadripartite structure. It featured a large single copy (LSC) region spanning from 89,158 to 89,336 bp, a small single copy (SSC) region ranging between 19,070 and 19,073 bp, and a pair of inverted repeat (IR) regions (IRa and IRb) spanning from 26,292 to 26,353 bp. The average GC content across the entire chloroplast genome was 37.4%, with specific percentages of 43% in the IR regions, 35.33% in the LSC region, and 31.67% in the SSC region. No significant differences were observed between the ten *V. vinifera* chloroplast genomes. Additionally, all chloroplast genomes shared highly similar gene contents, encompassing a total of 130 genes. These included 85 protein-coding genes, 37 tRNA genes, and 8 rRNA genes (Figure 1). Among these, 16 genes were found to be duplicated, comprising 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 (*rnn16*, *rnn23*, *rnn4.5*, and *rnn5*). Of these genes, 45 were associated with photosynthesis, while 74 were involved in various chloroplast transcription and translation activities. Furthermore, 11 protein-coding genes and 8 tRNA genes contained one intron each, whereas the *clpP1*, *pafI*, and *rps12B* genes were found to contain two introns each (Table 1 and Supplementary Table S1).

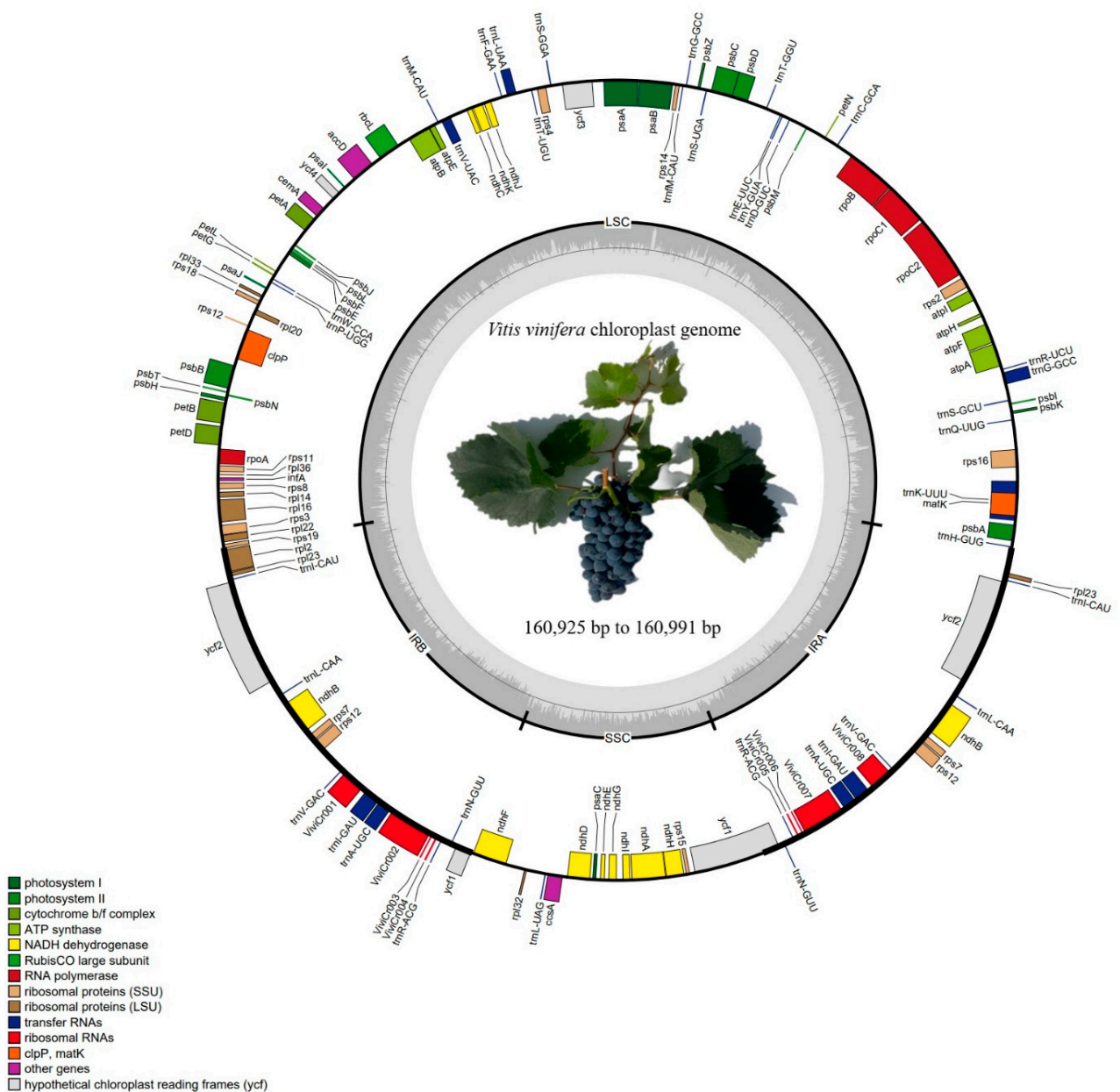


Figure 1. A circular chloroplast genome map of *V. vinifera* displaying genes located outside the circle that are transcribed in the clockwise direction, while those positioned inside the circle are transcribed counterclockwise. Genes are color-coded according to their functional groups. The dark gray region in the inner circle represents the GC content.

Table 1. Annotated genes and their classification in the chloroplast genomes of *V. vinifera* varieties.

Category	Group	Genes
Photosynthesis	Photosystem I	<i>psaA, psbA, psbC, psal, psaj</i>
	Photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbT, psbZ, psbN</i>
	ATP synthase	<i>atpA, atpB, atpE, atpF *, atpH, atpI</i>
	Cytochrome b/f complex	<i>petA, petB *, petD *, petN, petL, petG</i>
	Cytochrome C synthesis	<i>ccsA</i>
	RubisCO	<i>rbcl</i>
	NADPH dehydrogenase	<i>ndhA *, ndhB (×2) *, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>

Table 1. Cont.

Category	Group	Genes
	RNA polymerase	<i>rpoA, rpoB, rpoC2, rpoC1</i> *
	Ribosomal proteins	<i>rps2</i> (×2), <i>rps3, rps4, rps7</i> (×2), <i>rps8, rps11, rps12</i> (×2) *, <i>rps14, rps15, rps16</i> *, <i>rps18, rps19, rpl2</i> *, <i>rpl14, rpl16</i> *, <i>rpl20, rpl22, rpl23</i> (×2), <i>rpl32, rpl33, rpl36</i>
	Translation initiation factor	<i>infA</i>
Self-replication	Ribosomal RNA	<i>rrn16</i> (×2), <i>rrn23</i> (×2), <i>rrn4.5</i> (×2), <i>rrn5</i> (×2) <i>trna-UGC</i> (×2) *, <i>trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnFM-CAU, trnG-GCC, trnG-UCC</i> *, <i>trnH-GUG, trnI-CAU</i> (×2), <i>trnI-GAU</i> * (×2), <i>trnK-UUU</i> *, <i>trnL-CAA</i> (×2), <i>trnL-UAA</i> *, <i>trnL-UAG, trnM-CAU, trnN-GUU</i> (×2), <i>trnP-UGG, trnQ-UUG, trnR-ACG</i> (×2), <i>trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC</i> (×2), <i>trnV-UAC</i> *, <i>trnW-CCA, trnY-GUA</i>
	Maturase	<i>matK</i>
	Envelope membrane protein	<i>cemA</i>
Other genes	Acetyl-CoA carboxylase	<i>accD</i>
	Proteolysis	<i>ClpP</i> **
	Conserved ORFs	<i>ycf1</i> (×2), <i>ycf2</i> (×2), <i>ycf3, ycf4</i>

*: Gene containing a single intron; **: Gene containing two introns; (×2): Number of genes with multiple copies.

2.2. IR Boundary Analysis

The genome structure of *V. vinifera*, including the number and order of genes, exhibited a high degree of conservation across different samples. Nevertheless, variations were observed, particularly at the boundaries between the large single copy (LSC), inverted repeat (IR), and small single copy (SSC) regions. Expansion or contraction of these boundaries within the chloroplast genome influences its overall size and is a key mechanism driving size differences among genomes [28]. To investigate these structural variations, we utilized the initial *V. vinifera* reference (NC_007957.1) as a control. Our analysis focused on comparing the boundaries of the IR, LSC, and SSC regions across different varieties. Figure 2 illustrates some of these variations; yet, upon closer inspection, significant differences were not observed in these boundaries. Notably, at the LSC/IRb boundary, a fragment of the *rps19* pseudogene was identified in one set of varieties, spanning 46 bp within the IRb region. In contrast, in another set, this gene was entirely contained within the LSC region, approximately 17–18 bp away from the IRb boundary. The *ycf1* gene was consistently present across all five chloroplast genomes at the SSC/IRa boundary, without variation. This region, SSC/IRa, was found to be the most conserved, with the boundary consistently located within the 5681 bp-long *ycf1* gene in all samples. Furthermore, the IRa/SSC boundary in all examined samples was found to span across the *trnH* and *rps1* genes, exhibiting minor positional shifts from one variety to another.

2.3. Genomic Sequence Divergence

A sequence alignment analysis of the chloroplast genomes of the ten cultivars included in this study was performed using the chloroplast genome of *V. vinifera* (NC_007957.1) as a reference. The analysis demonstrated that the sequences were aligned in a consistent order and exhibited a high degree of conservation among the cultivars. We found that the non-coding regions showed greater divergence compared to protein-coding regions, indicating a more dynamic evolutionary history in these segments. Of the four major components of the chloroplast genome, the large single copy (LSC) region showed the most

significant variability, whereas the inverted repeat (IR) region exhibited the least (Figure 3). This data suggests that the LSC region may be subject to evolutionary pressures that drive change, while the IR region remains stable and conserved.

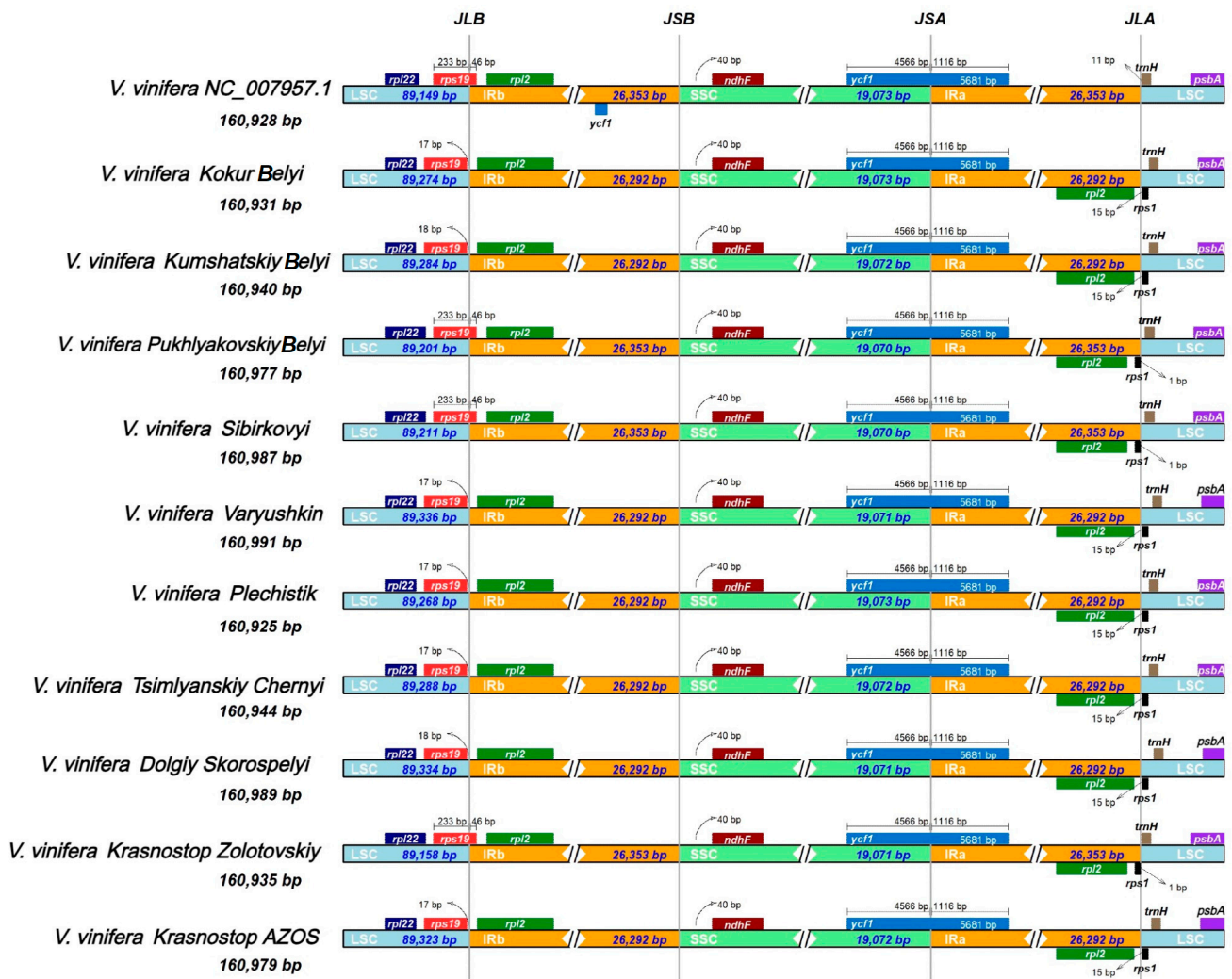


Figure 2. Comparison of the borders of the large single copy (LSC), small single copy (SSC), and inverted repeat (IR) regions in the chloroplast genomes of eleven *V. vinifera* varieties.

Furthermore, we evaluated nucleotide diversity (P_i) values through the analysis of 600 bp segments of the chloroplast genome, which enabled us to identify specific regions of sequence divergence. The results indicated a range of P_i values from 0 to 0.0038 across the ten genomes studied, highlighting the extent of genetic variation present. Notably, we identified five regions with particularly high variability ($P_i > 0.002$): the *trnS(GCU)*, *trnC(GCA)-petN*, *psbZ*, and *rpl20* genes, all located within the LSC region, as well as the *psaC* gene situated in the small single copy (SSC) region (Figure 4). These findings point to specific loci that may be of interest for further investigations into the evolutionary dynamics and adaptation mechanisms of these cultivars. By focusing on these variable regions, future studies could enhance our understanding of the genetic diversity within chloroplast genomes and their implications for plant biology and breeding strategies.

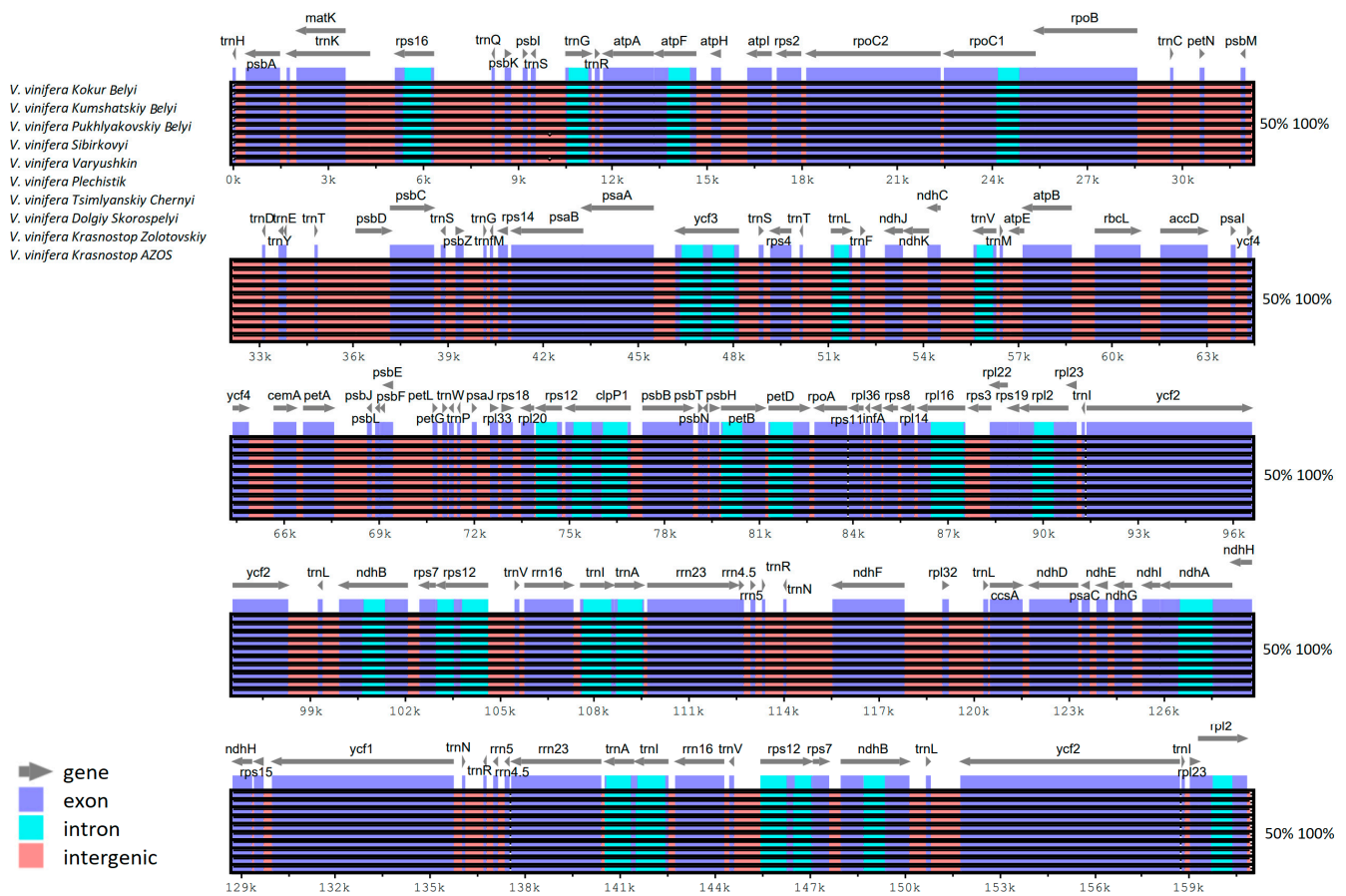


Figure 3. The sequence alignment of ten *V. vinifera* genomes in MVISTA is illustrated. The gray arrows above the alignment indicate the transcription directions of the genes. The y-axis displays the identity percentage, which ranges from 50% to 100% for each comparison. Genome regions are color-coded as exon, intron, and intergenic regions.

2.4. Identification of Long Repeats

Repetitive sequences are crucial for elucidating phylogenetic relationships among species and play a significant role in genome rearrangement [22]. In our study, we identified several types of repeats—forward (F), reverse (R), complementary (C), and palindromic (P)—in the chloroplast genomes of ten *V. vinifera* cultivars. In total, we detected 99 repeated sequences across these genomes, excepting the Sibirkoviy and Dolgiy Skorospelyi cultivars, which each contained 98 repeats (Figure 5 and Supplementary Table S2). The distribution of repeats exhibited considerable variation among the cultivars. The number of direct repeats ranged from 36 to 40, reverse repeats from 14 to 18, complementary repeats from 4 to 6, and palindromic repeats from 39 to 42. The Krasnostop AZOS variety exhibited the highest number of palindromic repeats at 42, while the counts of other types of repeat sequences were relatively consistent across the majority of *V. vinifera* cultivars. A significant finding of our analysis was that many of these repetitive sequences were predominantly located within non-coding regions, particularly in intergenic and intronic regions. This distribution suggests that these regions may play a critical role in the structural dynamics and regulatory mechanisms of the chloroplast genome [23]. Repetitive sequences in non-coding regions may offer insights into evolutionary processes, as these areas are more prone to variations that can drive genome rearrangement [24,25]. Further research on the functional implications of these repeats might illuminate their contributions to genetic diversity, adaptation, and potential breeding strategies in grapevine cultivars. Such investigations could enhance our understanding of the evolutionary significance of repetitive sequences in plant genomes as a whole.

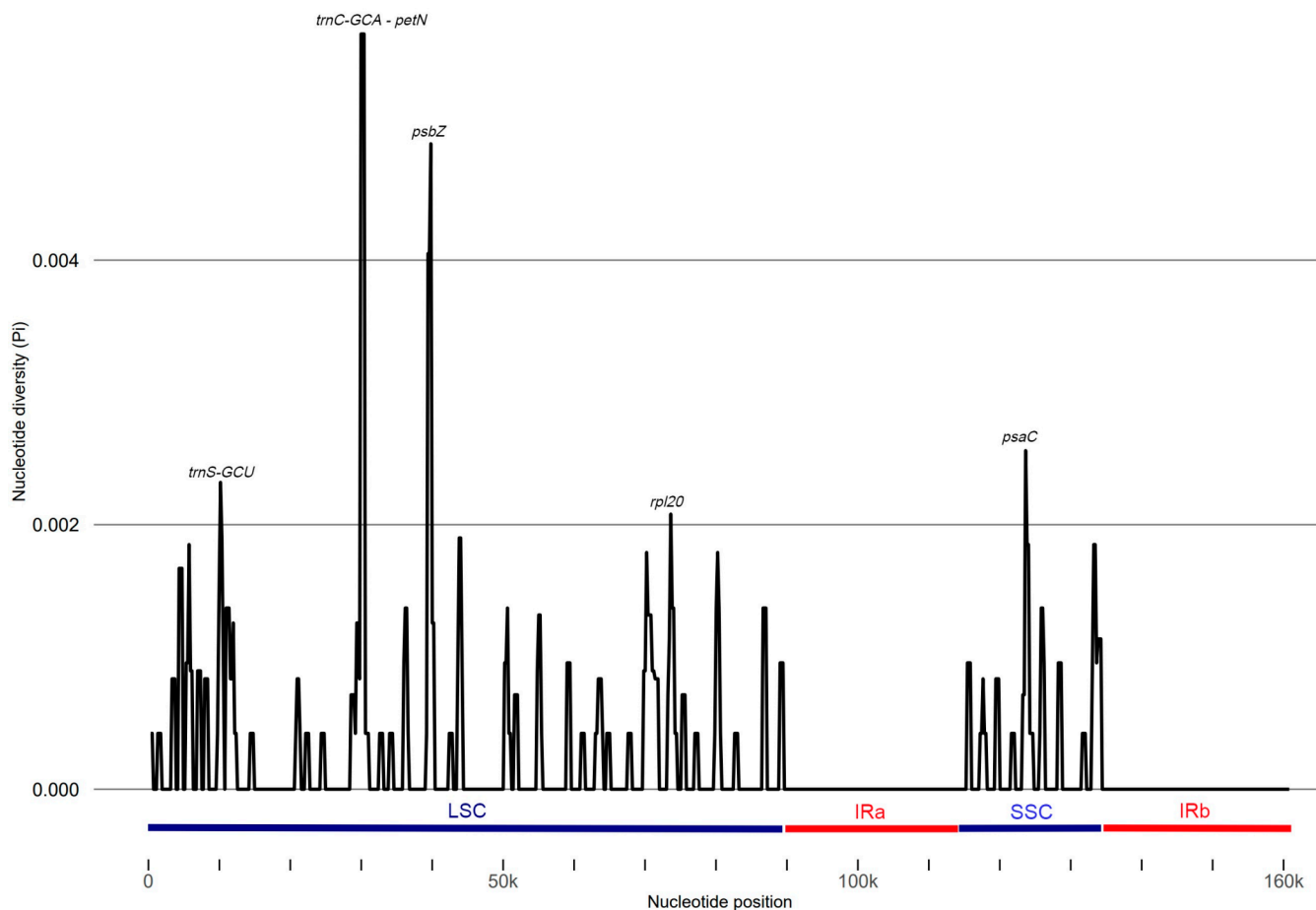


Figure 4. Nucleotide diversity (Pi) analysis for chloroplast genomes from the *V. vinifera* varieties with window length: 600 bp; step size: 200 bp.

2.5. Identification Simple Sequence Repeats

Simple sequence repeats (SSRs), commonly referred to as microsatellites, are characterized by relatively high mutation rates and copy number polymorphisms. This makes them invaluable as molecular markers for studies of genetic diversity, polymorphism, evolutionary dynamics, and plant breeding [29]. Furthermore, the study of complex SSRs can provide insights into the evolutionary processes that underpin microsatellite formation and variation [30]. In our study of the chloroplast genomes of grape varieties, we found that the SSR content varied across different cultivars. Specifically, the number of SSRs ranged from 58 in the varieties Kumshatskiy Belyi, Varyushkin, and Tsimlyanskiy Chernyi to 61 SSRs in the varieties Dolgiy Skorospelyi and Sibirkovyi. Among these, the percentage of mononucleotide repeats was highest in the Sibirkovyi variety, which exhibited 57 such repeats. In contrast, the number of dinucleotide and trinucleotide repeats was consistently recorded as two across most studied varieties, with the exception of the Dolgiy Skorospelyi variety, which had a third trinucleotide repeat (Figure 6 and Supplementary Table S3). The analysis revealed that the majority of mononucleotide and dinucleotide repeats were composed of A/T and AT/TA combinations, with other repeat types also predominantly featuring A and T as the major repeat units. In comparison, the occurrences of C and G were quite rare, indicating a bias in nucleotide composition for these SSRs. Upon examination of the chloroplast regions, it was observed that the highest SSR content was present in the large single copy (LSC) region, while the inverted repeat (IR) region exhibited the lowest SSR density. This distribution underscores the significance of the LSC region in contributing to genetic variability, which may have implications for understanding the adaptive potential and evolutionary pathway of grape varieties.

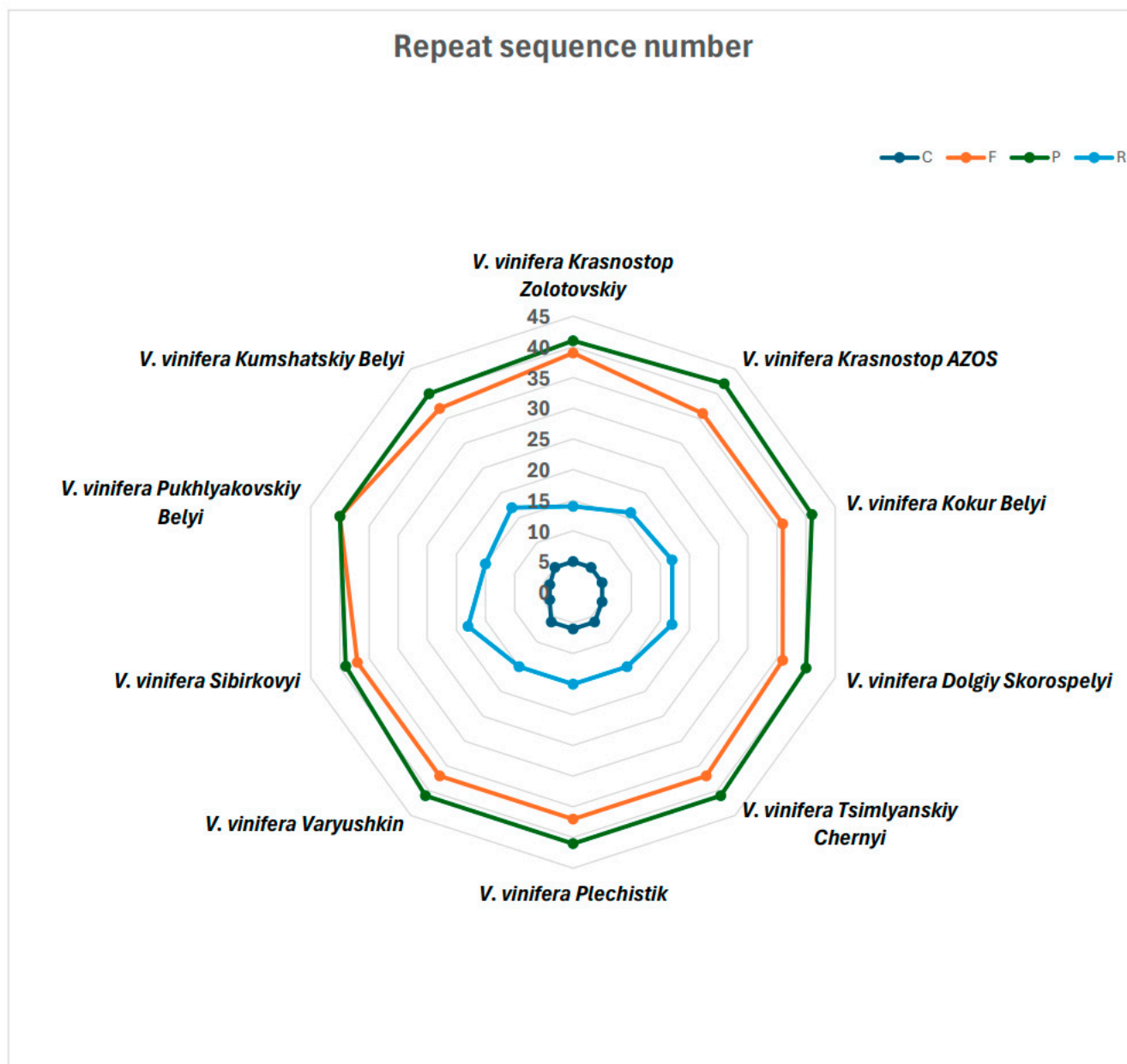


Figure 5. The number of long repeat sequences in the ten chloroplast genomes of *V. vinifera* (F: forward repeats; R: reverse repeats; P: palindromic repeats; C: complementary repeats).

2.6. Selection Pressure Analysis

The ratios of nonsynonymous (K_a) to synonymous (K_s) substitutions for common protein-coding genes (PCGs) were calculated across ten grape genomes, with *V. vinifera* (NC_007957.1) and *V. vinifera* subsp. *sylvestris* (LC523806.1) as reference sequences (Supplementary Table S4). Out of the 85 common genes analyzed, we encountered challenges in determining the K_a/K_s ratios for many due to a lack of identical sequence variations or the absence of nonsynonymous or synonymous nucleotide substitutions. Additionally, in cases where calculations were possible, many did not yield significant results (p -value > 0.05). Interestingly, the majority of K_a/K_s ratios were found to be less than one, suggesting that most genes were under negative selection, indicating a preference for conserving protein function over time. However, one gene, *rpoC2*, exhibited potential signs of positive selection across all grape varieties, with a K_a/K_s ratio approaching 1. This could imply

adaptive evolution in this gene, which may be associated with specific functions that confer advantages in certain environmental conditions or developmental processes [31,32].

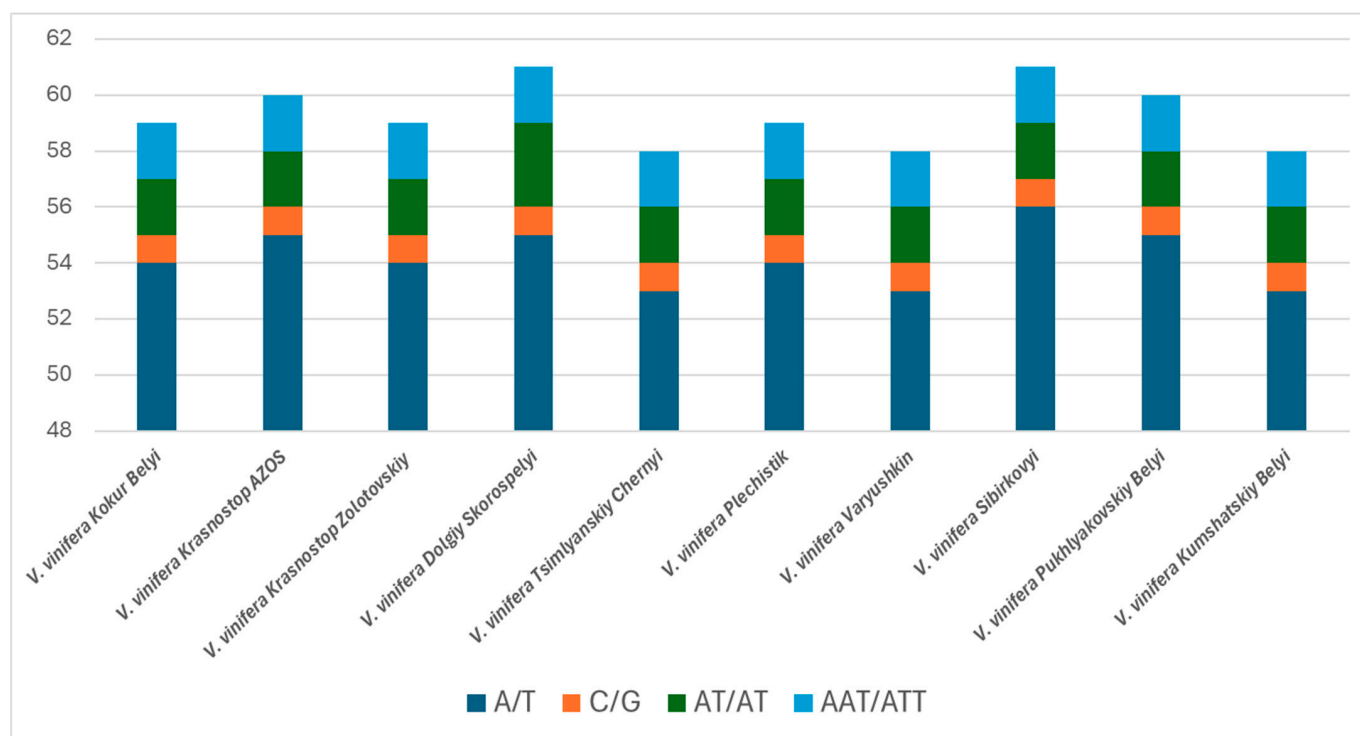


Figure 6. Number and types of SSRs detected in the ten chloroplast genomes of *V. vinifera*.

2.7. Codon Usage Bias

We carried out an analysis of codon usage bias within the chloroplast genomes of *V. vinifera*, focusing on 85 protein-coding genes in order to gain a deeper understanding of the patterns in amino acid representation. Codons with a relative synonymous codon usage (RSCU) value greater than 1 were identified as being preferentially utilized for encoding specific amino acids, providing insights into genomic codon selection preferences [33]. This study examined the relative usage rates of codons across all protein-coding genes in the chloroplast genomes of ten distinct *V. vinifera* varieties, calculating the synonymous codon usage rates for comparison. The analysis revealed that all varieties contained a total of 61 different codons utilized to encode 22 amino acids (Supplementary Table S5). Interestingly, seven varieties shared the same total number of codons (23,273), while the varieties Tsimlyanskiy Chernyi and Plechistik had slightly fewer at 22,597, and the Sibirkovyi variety exhibited the smallest number of codons, totaling 23,184. Among the amino acids analyzed, leucine (*Leu*) emerged as the most frequently represented, accounting for 10.31–10.41% of codons, while cysteine (*Cys*) had the lowest representation at 1.11–1.13%. The AUU codon, which encodes isoleucine (*Ile*), was found to be the most common overall, occurring between 951 and 980 times across the genomes. Conversely, the UGC codon for cysteine (*Cys*) was the least prevalent, with only 71 to 74 total occurrences. In total, we identified 31 codon usage preferences (RSCU > 1) across the chloroplast genomes of all ten *V. vinifera* varieties, indicating significant bias in codon selection. The codons showcasing the highest and lowest RSCU values were UUA, encoding leucine, and GGC, which encodes glycine.

2.8. Phylogenetic Analysis

To determine the phylogenetic position of autochthonous *V. vinifera* cultivars from the Don Valley, an ML (maximum likelihood) tree was constructed using the protein-coding genes of the *Vitis* chloroplast genomes obtained in the present study, as well as the other

15 members of the *Vitis* family downloaded from the NCBI database with *Vitis rotundifolia* as an outgroup. Phylogenetic analysis showed that, in addition to the variety *V. vinifera* Krasnostop Zolotovskiy, the varieties Tsimlyanskiy Chernyi, Plechistik, and Krasnostop AZOS also did not have direct genetic relationships with other Caucasian and European varieties of *V. vinifera* (Figure 7), as was shown earlier [34]. Instead, they formed a separate cluster with other autochthonous varieties in the genetic dendrogram of *V. vinifera*, which sharply contradicts earlier theories that suggested they originated from Eastern European varieties introduced into Russia. At the same time, the European and Caucasian varieties formed two separate groups with equally strong support (BS = 100). The Caucasian group included the majority of the autochthonous varieties analyzed in this study (Varyushkin, Dolgiy Skorospelyi, Puklyakovskiy Belyi, Sibirkovyi, and Kumshatskiy Belyi), while the European cluster included only one variety—Kokur Belyi. The aforementioned results demonstrate that chloroplast genome sequence analysis can offer valuable insights into the genetic background of *V. vinifera* varieties [35]. These can be used to aid in breeding and provide a molecular biological basis for cultivar identification.

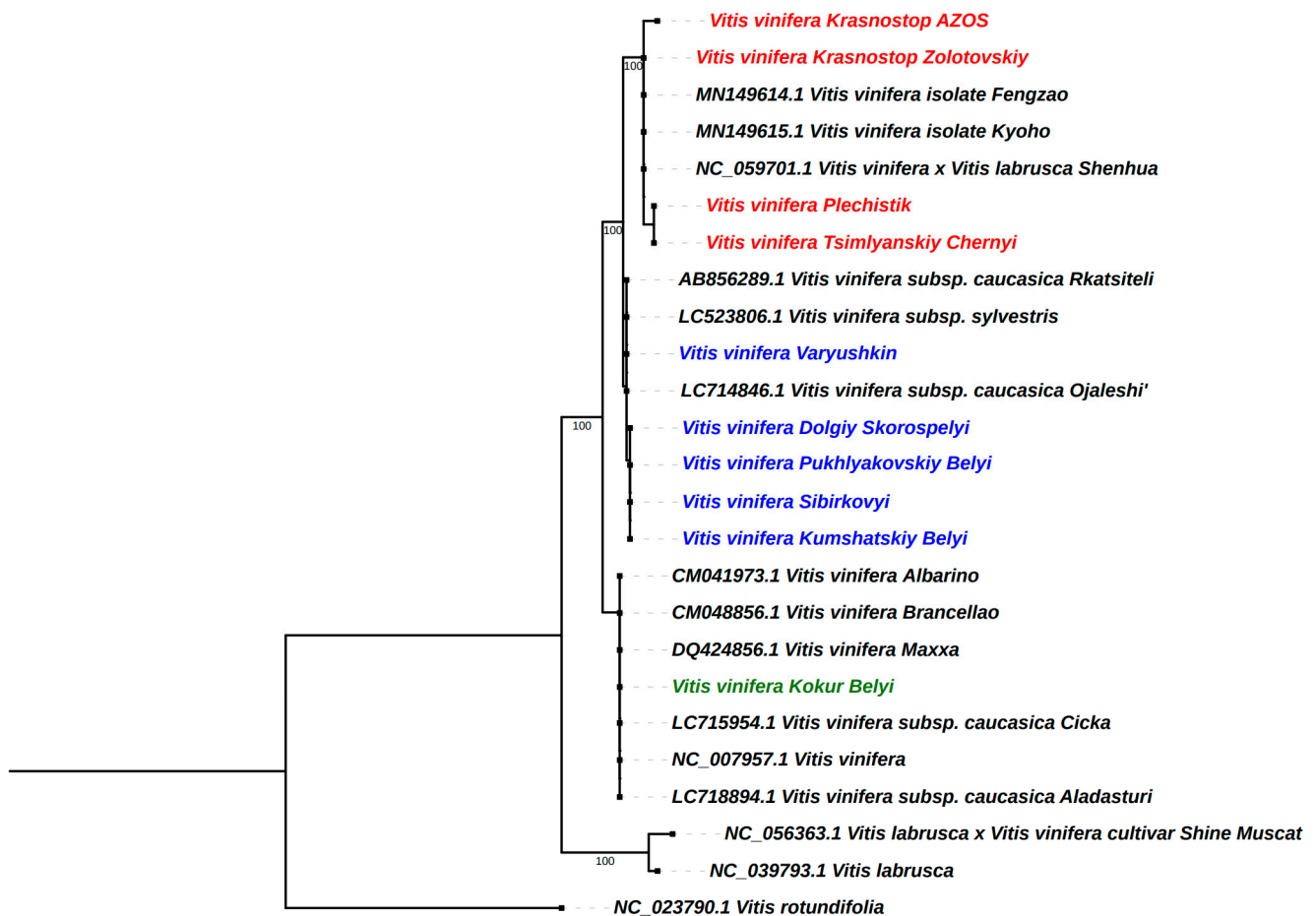


Figure 7. Phylogenetic reconstruction of *Vitis* using maximum likelihood (ML) methods based on PCGs. The values above the nodes are bootstrap support (BS) values.

3. Discussion

The present study involved the sequencing, assembly, and annotation of eight complete chloroplast genomes of autochthonous grape varieties from the Don Valley and one from Crimea. In order to gain a comprehensive understanding of the chloroplast genomes of grape varieties, we conducted a detailed comparison of the Krasnostop Zolotovskiy genome, which was previously analyzed in our laboratory [34], with those of other varieties. Our analysis encompassed the content of genes, long repeats, microsatellites, and stop

codons. Additionally, we examined regions exhibiting the highest degree of variability and searched for genes displaying evidence of positive selection. Nevertheless, a comparison of chloroplast genomes in different cultivars of the same species revealed a high degree of conservation.

The conservation of chloroplast genomes in different grape varieties (*V. vinifera*) is of significant importance to our understanding of their genetic structures and evolutionary relationships [13]. The chloroplast genomes of grapes, like those of most higher plants [36], demonstrate a significant degree of homology between varieties, which allows them to be used for the study of phylogenetic relationships and the identification of varieties [37]. In this study, we aimed to clarify the phylogenetic relationships of autochthonous *V. vinifera* cultivars from the Don Valley by constructing a maximum likelihood (ML) tree based on protein-coding genes from *Vitis* chloroplast genomes. Our analysis showed that some local cultivars, including Tsimlyanskiy Chernyi, Plechistik, and Krasnostop AZOS, along with Krasnostop Zolotovskiy, do not have direct genetic relationships with other Caucasian and European varieties of *V. vinifera*. This finding contradicts the prevailing theory that these varieties originated from Eastern European varieties introduced to Russia [35]. The formation of a discrete cluster of these autochthonous varieties indicates the potential for disparate evolutionary trajectories and adaptation processes due to the distinctive ecological and climatic conditions of the Don Valley region [1]. This discrepancy emphasizes the significance of local cultivation practices and the particular ecological niches that these varieties occupy, thereby enabling them to evolve distinctive genetic attributes over time. The results also demonstrate a clear separation between the European and Caucasian groups within the genetic dendrogram. The Caucasian cluster is represented mainly by the varieties studied in the work, while the European cluster is represented by one variety, Kokur Belyi. This difference raises questions about the historical and geographic factors that contributed to the genetic diversity observed among *V. vinifera* cultivars. This may also challenge the assumption that many autochthonous varieties share a common origin with their European counterparts.

The grape chloroplast genome features a circular molecular structure containing dozens of genes responsible for photosynthesis and other metabolic functions. Throughout evolution, most genes related to essential chloroplast functions have remained largely unchanged. This indicates a strong selective pressure for their conservation [38]. Key genes, such as those encoding photosystem subunits and enzymes involved in chlorophyll synthesis, maintain conserved sequences, underscoring their crucial role in plant survival [39]. The results of this study emphasize the significance of simple sequence repeats (SSRs) in the genetic analysis of grapes, which may result in future studies using these markers for genetic mapping, cultivar identification, and the improvement of breeding programs [40]. Our results indicate marked differences in SSR content between different cultivars, with the number of SSRs varying from 58 to 61 among the cultivars analyzed. This variability underscores the potential of SSRs for use in cultivar discrimination and understanding the genetic structure of grape populations. Investigating the functional consequences of SSR variation in grape chloroplast genomes enhances our understanding of evolutionary processes. It also helps to determine the mechanisms of diversification of this economically important plant [41].

Additionally, our results revealed differences in codon usage among *V. vinifera* cultivars, illuminating their evolutionary adaptations [42]. Understanding these codon preferences is crucial for developing genetic markers for breeding and conservation, as well as for studying the functional roles of proteins in grapes.

Our research findings have significant implications for grapevine breeding and conservation strategies. The detailed genetic characterization of these varieties may result in the discovery of unique traits that are valuable for breeding programs focused on disease resistance, climate adaptability, and general agronomic performance [43]. By analyzing chloroplast genome sequences, researchers and breeders can gain a deeper understanding

of the genetic basis of *V. vinifera*, leading to more informed selection of parental varieties and improved cultivar development.

4. Materials and Methods

4.1. Plant Materials, DNA Extraction, and Sequencing

In this study, ten samples of *Vitis vinifera* were collected from different places in southern Russia (Supplementary Table S1). Fresh and healthy leaves of these ten samples were collected. The DNA extraction protocol used in this research was adjusted from that of Sandra Lo Piccolo (2012) [44]. Modifications included performing all centrifugation procedures at $3500 \times g$ and 4°C [8,45]. The centrifugation time was 15 min and 45 min for DNA extraction with chloroform-isoamyl alcohol and isopropanol precipitation of DNA, respectively. 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.

4.2. Genome Assembly and Annotation

Fastp v0.23.4 [35] was employed to eliminate adapter-containing sequences and low-quality reads from the data. This process resulted in clean read yields ranging from 56.20 to 129.32 GB. The chloroplast genomes were then assembled using GetOrganelle software (version 1.7.7.0) [36] with default settings. Following assembly, the contig underwent correction with Pilon [37] and validation through short-read mapping using Bowtie2 (version 2.4.4) [38]. The annotation of the assembled chloroplast genomes was conducted using GeSeq v.1 software [39]. For the identification of tRNA genes, we utilized tRNAscanSE 2.0.12 [40] with default parameters. The rRNA components of the chloroplast genome were annotated using BLASTN software (version 2.12.0) [41]. The circular map of the chloroplast genome was visualized using OGDRAW 1.3.1 [42]. All read data generated during this study have been documented and deposited in the NCBI Sequence Read Archive (SRA) and can be accessed under the respective accession number PRJNA1011053.

4.3. Comparative Genome Analysis

The nine complete chloroplast genomes of *V. vinifera* in this study and another variety, *V. vinifera* Krasnostop Zolotovskiy, were compared using the MVISTA [46] program with the shuffle-LAGAN model [47], with *V. vinifera* (NC_007957.1) as the reference. The IRSCOPE program [48] was applied to analyze the LSC, IR, and LSC boundary locations in eleven complete *Vitis* chloroplast genomes. Sliding window analysis was conducted to determine the nucleotide variability (Pi) of the complete chloroplast genome using DnaSP v6 [49] after sequence alignment with MAFFT v7.427 [50]. The sliding window length was set as 600 bp, with a step size of 200 bp. In addition, DnaSP v6 was adopted to calculate insertions and deletions (InDels) and Pi for highly variable regions.

4.4. Analysis of the Cp Genome for Repetitive Sequences

Simple sequence repeats (SSRs) in the chloroplast genomes of ten *Vitis* varieties were analyzed using MISA v2.1 software [51]. The minimum repeat units were configured as follows: 10 for mononucleotide SSRs, 5 for dinucleotide SSRs, 4 for trinucleotide SSRs, and 3 for tetranucleotide, pentanucleotide, and hexanucleotide SSRs. Additionally, REPuter v1 software [52] was employed to identify different types of repeats, including forward (F), reverse (R), palindrome (P), and complementary (C), with criteria set for a minimum repeat size of 30 bp and a Hamming distance of 3.

4.5. Codon Preference Analysis

Codon W v1.4.4 software [53] was utilized for codon analysis and to determine relative synonymous codon usage (RSCU) values for comparative mapping. RSCU reflects the likelihood of employing synonymous codons to encode a particular amino acid: an RSCU value greater than 1 indicates that the codon is frequently used; an RSCU value equal to 1 suggests no bias in codon usage; and an RSCU value of less than 1 indicates that the codon is rarely used.

4.6. Selective Analysis

To evaluate selective pressure on the *Vitis* varieties, we analyzed the Ka/Ks ratio (where Ka represents the nonsynonymous substitution rate and Ks denotes the synonymous substitution rate) in a study involving ten grapevine varieties. Using PhyloSuite v1.2.3 [54], we extracted 85 common protein-coding genes (PCGs) and translated them into their corresponding amino acids. We then utilized KaKs_calculator 3.0 [55] with ParaAT v2.0 [56] to automatically generate intermediate files and compute the Ka/Ks values, discarding results with excessively high Na values, which were likely due to an extremely low synonymous substitution ratio. A Ka/Ks ratio greater than 1 indicated that the gene pair was under positive selection, while a ratio lower than 1 suggested purifying selection.

4.7. Phylogenetic Analysis

For the phylogenetic analysis, the complete chloroplast genomes of 15 samples within the *Vitis* family, available by June 2024, were downloaded from the National Center for Biotechnology Information (NCBI). *Vitis rotundifolia* were chosen as the outgroups. Common protein-coding genes (PCGs) were extracted and aligned using MAFFT v7.427 [50] using the iterative method (G-INS-i) using default parameter settings. To construct the maximum likelihood (ML) phylogenetic tree, the RAxML algorithm [57] was employed and bootstrap values were obtained from 1000 replicates. The resulting tree was visualized using the iTOL service [58].

5. Conclusions

For the first time, a comparative analysis of chloroplasts in grape genomes was carried out. But, since the chloroplast genomes in open databases today are quite small, we made a significant contribution to the further study of chloroplasts as markers for studying the adaptive and evolutionary mechanisms of this genus. Studying the chloroplast genomes of grapes can help biodiversity conservation programs by tracking genetic resources and risks to extinction in certain varieties. Thus, chloroplast genomes are an important object of scientific research due to their importance in photosynthesis, heredity, environmental adaptation, and biotechnology.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25189928/s1>.

Author Contributions: Conceptualization, M.V.P. and D.Y.F.; methodology, S.V.T. and F.S.S.; collection and validation of specimens, D.Y.F.; data curation, K.O.P. and F.S.S.; validation, D.Y.F. and F.S.S.; formal analysis, F.S.S. and K.O.P.; writing—original draft preparation D.Y.F. and F.S.S.; writing—review and editing, D.Y.F., F.S.S. and S.V.T.; visualization, F.S.S.; project administration, S.V.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work was carried out within the state assignment of the NRC “Kurchatov institute” and supported by the Ministry of Science and Higher Education of the Russian Federation, grant #075-15-2019-1659.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The chloroplast sequences of *V. vinifera* varieties were deposited into the National Center for Biotechnology Information (NCBI) repository, and the accession number is PRJNA1011053.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Petrova, K.O.; Fedosov, D.Y.; Korzhenkov, A.A.; Toshchakov, S.V.; Patrushev, M.V.; Chernikov, V.Y. On the gene pool of don valley grape varieties. *Nanobiotechnol. Rep.* **2022**, *17*, 722–730. [[CrossRef](#)]
2. Mihaljević, M.Ž.; Maletić, E.; Preiner, D.; Zdunić, G.; Bubola, M.; Zyprian, E.; Pejić, I. Genetic diversity, population structure, and parentage analysis of Croatian grapevine germplasm. *Genes* **2020**, *11*, 737. [[CrossRef](#)] [[PubMed](#)]
3. Manatskov, A.G.; Petrov, V.S.; Naumova, L.G. Autochthonous grape varieties in the Lower Don Region. *Sadovod. Vinograd.* **2024**, *2*, 22–29. [[CrossRef](#)]
4. Kraysvetniy, M.I. The nart epos and the don winemaking. *Russ. Vine* **2018**, *8*, 116–131. [[CrossRef](#)]
5. Anderson, K. JANCIS ROBINSON, JULIA HARDING and JOSÉ VOUILLAMOZ: Wine Grapes: A Complete Guide to 1368 Vine Varieties, including their Origins and Flavours. Ecco (Harper Collins), New York, October 2012, xxxvii + 1242 pp., ISBN 978-0062206367 (hardback), US\$175. *J. Wine Econ.* **2013**, *8*, 106–109. [[CrossRef](#)]
6. Pankin, M.I.; Petrov, V.S.; Lukianova, A.A.; Il'nitskaya, E.T.; Nikulushkina, G.E.; Kovalenko, A.G.; Bolshakov, V.A. The anapa ampelographic collection is the largest center of vine gene pool accumulation and research in Russia. *Vavilov J. Genet. Breed.* **2018**, *22*, 54–59. [[CrossRef](#)]
7. Rybalko, E.; Ostroukhova, E.; Peskova, I.; Romanov, A.; Boyko, V. Crimean autochthonous grape varieties as a factor of high-quality winemaking in a changing climate. *BIO Web Conf.* **2022**, *53*, 01001. [[CrossRef](#)]
8. Volynkin, V.; Polulyakh, A.; Levchenko, S.; Vasylyk, I.; Likhovskoy, V. Aspects of the particular genetics of grapes prolonged for all horticulture crops. In *Horticultural Crops*; Kossi Baimey, H., Hamamouch, N., Adjiguita Kolombia, Y., Eds.; IntechOpen: London, UK, 2020; ISBN 978-1-83880-421-3.
9. Il'nitskaya, E.T.; Tokmakov, S.V.; Suprun, I.I.; Naumova, L.G.; Ganich, V.A. Genetic similarity of the autochthonous grapevine varieties from don region revealed by SSR-analysis and main leaf ampelographic traits. *Sel'skokhozyaistvennaya Biol.* **2016**, *51*, 60–67. [[CrossRef](#)]
10. Ganich, V.A.; Naumova, L.G. Kumshatsky Belyy—Perspective aboriginal Don grapevine variety. *Bull. KrasGAU* **2021**, *12*, 11–16. [[CrossRef](#)]
11. Wolfe, K.H.; Li, W.H.; Sharp, P.M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 9054–9058. [[CrossRef](#)]
12. Neuhaus, H.E.; Emes, M.J. Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2000**, *51*, 111–140. [[CrossRef](#)] [[PubMed](#)]
13. Robbins, E.H.J.; Kelly, S. The evolutionary constraints on angiosperm chloroplast adaptation. *Genome Biol. Evol.* **2023**, *15*, evad101. [[CrossRef](#)] [[PubMed](#)]
14. Sharko, F.; Gladysheva-Azgari, M.; Tsygankova, S.; Mitrofanova, I.; Boulygina, E.; Slobodova, N.; Smykov, A.; Rastorguev, S.; Nedoluzhko, A. The complete chloroplast genome sequence of cultivated *Prunus persica* cv. "Sovetskiy". *Mitochondrial DNA Part B Resour.* **2021**, *6*, 2882–2883. [[CrossRef](#)]
15. Gao, H.; McCormick, A.J.; Roston, R.L.; Lu, Y. Editorial: Structure and function of chloroplasts, Volume III. *Front. Plant Sci.* **2023**, *14*, 1145680. [[CrossRef](#)] [[PubMed](#)]
16. Bock, R. Structure, function, and inheritance of plastid genomes. In *Cell and Molecular Biology of Plastids*; Bock, R., Ed.; Topics in Current Genetics; Springer: Berlin/Heidelberg, Germany, 2007; Volume 19, pp. 29–63. ISBN 978-3-540-75375-9.
17. Ohyama, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shiki, Y.; Takeuchi, M.; Chang, Z.; et al. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **1986**, *322*, 572–574. [[CrossRef](#)]
18. Dong, W.; Liu, J.; Yu, J.; Wang, L.; Zhou, S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS ONE* **2012**, *7*, e35071. [[CrossRef](#)]
19. Daniell, H.; Lin, C.-S.; Yu, M.; Chang, W.-J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [[CrossRef](#)]
20. Jansen, R.K.; Ruhlman, T.A. Plastid Genomes of Seed Plants. In *Genomics of Chloroplasts and Mitochondria*; Bock, R., Knoop, V., Eds.; Advances in Photosynthesis and Respiration; Springer: Dordrecht, The Netherlands, 2012; Volume 35, pp. 103–126, ISBN 978-94-007-2919-3.
21. Herrmann, L.; Haase, I.; Blauhut, M.; Barz, N.; Fischer, M. DNA-based differentiation of the Ecuadorian cocoa types CCN-51 and Arriba based on sequence differences in the chloroplast genome. *J. Agric. Food Chem.* **2014**, *62*, 12118–12127. [[CrossRef](#)] [[PubMed](#)]
22. Khan, A.L.; Asaf, S.; Lee, I.-J.; Al-Harrasi, A.; Al-Rawahi, A. First chloroplast genomics study of *Phoenix dactylifera* (var. Naghal and Khanezi): A comparative analysis. *PLoS ONE* **2018**, *13*, e0200104. [[CrossRef](#)]

23. Kahlau, S.; Aspinnall, S.; Gray, J.C.; Bock, R. Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *J. Mol. Evol.* **2006**, *63*, 194–207. [[CrossRef](#)]
24. Tang, J.; Xia, H.; Cao, M.; Zhang, X.; Zeng, W.; Hu, S.; Tong, W.; Wang, J.; Wang, J.; Yu, J.; et al. A comparison of rice chloroplast genomes. *Plant Physiol.* **2004**, *135*, 412–420. [[CrossRef](#)] [[PubMed](#)]
25. Mariotti, R.; Cultrera, N.G.M.; Díez, C.M.; Baldoni, L.; Rubini, A. Identification of new polymorphic regions and differentiation of cultivated olives (*Olea europaea* L.) through plastome sequence comparison. *BMC Plant Biol.* **2010**, *10*, 211. [[CrossRef](#)] [[PubMed](#)]
26. Wei, R.; Li, Q. The Complete Chloroplast Genome of Endangered Species *Stemona parviflora*: Insight into the Phylogenetic Relationship and Conservation Implications. *Genes* **2022**, *13*, 1361. [[CrossRef](#)]
27. Bock, R. Engineering plastid genomes: Methods, tools, and applications in basic research and biotechnology. *Annu. Rev. Plant Biol.* **2015**, *66*, 211–241. [[CrossRef](#)] [[PubMed](#)]
28. Huang, H.; Shi, C.; Liu, Y.; Mao, S.-Y.; Gao, L.-Z. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: Genome structure and phylogenetic relationships. *BMC Evol. Biol.* **2014**, *14*, 151. [[CrossRef](#)]
29. Li, B.; Lin, F.; Huang, P.; Guo, W.; Zheng, Y. Development of nuclear SSR and chloroplast genome markers in diverse *Liriodendron chinense* germplasm based on low-coverage whole genome sequencing. *Biol. Res.* **2020**, *53*, 21. [[CrossRef](#)]
30. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol.* **2005**, *23*, 48–55. [[CrossRef](#)] [[PubMed](#)]
31. Ren, T.; Li, Z.-X.; Xie, D.-F.; Gui, L.-J.; Peng, C.; Wen, J.; He, X.-J. Plastomes of eight *Ligusticum* species: Characterization, genome evolution, and phylogenetic relationships. *BMC Plant Biol.* **2020**, *20*, 519. [[CrossRef](#)]
32. Gladysheva-Azgari, M.; Sharko, F.; Slobodova, N.; Petrova, K.; Boulygina, E.; Tsygankova, S.; Mitrofanova, I. Comparative Analysis Revealed Intrageneric and Intraspecific Genomic Variation in Chloroplast Genomes of *Actinidia* spp. (*Actinidiaceae*, *Viridiplantae*). *Horticulturae* **2023**, *9*, 1175. [[CrossRef](#)]
33. Nie, L.; Cui, Y.; Wu, L.; Zhou, J.; Xu, Z.; Li, Y.; Li, X.; Wang, Y.; Yao, H. Gene Losses and Variations in Chloroplast Genome of Parasitic Plant *Macrosolen* and Phylogenetic Relationships within Santalales. *Int. J. Mol. Sci.* **2019**, *20*, 5812. [[CrossRef](#)]
34. Sharko, F.S.; Petrova, K.O.; Molchanov, N.V.; Patrushev, M.V.; Fedosov, D.Y.; Toshchakov, S.V. The complete chloroplast genome sequence of *Vitis vinifera* ‘Krasnostop Zolotovskiy’, an autochthonous variety of the Don Valley. *Turczaninowia* **2024**, *27*, 124–129. [[CrossRef](#)]
35. 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. SNP-Based Analysis Reveals Authenticity and Genetic Similarity of Russian Indigenous *V. vinifera* Grape Cultivars. *Plants* **2021**, *10*, 2696. [[CrossRef](#)] [[PubMed](#)]
36. Meger, J.; Ulaszewski, B.; Vendramin, G.G.; Burczyk, J. Using reduced representation libraries sequencing methods to identify cpDNA polymorphisms in European beech (*Fagus sylvatica* L.). *Tree Genet. Genomes* **2019**, *15*, 7. [[CrossRef](#)]
37. Somaratne, Y.; Guan, D.-L.; Wang, W.-Q.; Zhao, L.; Xu, S.-Q. The Complete Chloroplast Genomes of Two *Lespedeza* Species: Insights into Codon Usage Bias, RNA Editing Sites, and Phylogenetic Relationships in Desmodieae (Fabaceae: Papilionoideae). *Plants* **2019**, *9*, 51. [[CrossRef](#)]
38. Wicke, S.; Schneeweiss, G.M.; dePamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [[CrossRef](#)]
39. Jarvis, P.; López-Juez, E. Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 787–802. [[CrossRef](#)] [[PubMed](#)]
40. Ebert, D.; Peakall, R. Chloroplast simple sequence repeats (cpSSRs): Technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Mol. Ecol. Resour.* **2009**, *9*, 673–690. [[CrossRef](#)]
41. George, B.; Bhatt, B.S.; Awasthi, M.; George, B.; Singh, A.K. Comparative analysis of microsatellites in chloroplast genomes of lower and higher plants. *Curr. Genet.* **2015**, *61*, 665–677. [[CrossRef](#)]
42. Qin, Z.; Cai, Z.; Xia, G.; Wang, M. Synonymous codon usage bias is correlative to intron number and shows disequilibrium among exons in plants. *BMC Genom.* **2013**, *14*, 56. [[CrossRef](#)]
43. Jiao, L.; Lu, Y.; He, T.; Li, J.; Yin, Y. A strategy for developing high-resolution DNA barcodes for species discrimination of wood specimens using the complete chloroplast genome of three *Pterocarpus* species. *Planta* **2019**, *250*, 95–104. [[CrossRef](#)]
44. Lo Piccolo, S.; Alfonso, G.; Conigliaro, G.; Moschetti, G.; Burruano, S.; Barone, A. A simple and rapid DNA extraction method from leaves of grapevine suitable for polymerase chain reaction analysis. *Afr. J. Biotechnol.* **2012**, *11*, 10305. [[CrossRef](#)]
45. Gladysheva-Azgari, M.; Petrova, K.; Tsygankova, S.; Mitrofanova, I.; Smykov, A.; Boulygina, E.; Slobodova, N.; Rastorguev, S.; Sharko, F. A de novo genome assembly of cultivated *Prunus persica* cv. “Sovetskiy”. *PLoS ONE* **2022**, *17*, e0269284. [[CrossRef](#)] [[PubMed](#)]
46. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, W273–W279. [[CrossRef](#)] [[PubMed](#)]
47. Brudno, M.; Do, C.B.; Cooper, G.M.; Kim, M.F.; Davydov, E.; NISC Comparative Sequencing Program; Green, E.D.; Sidow, A.; Batzoglou, S. LAGAN and Multi-LAGAN: Efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **2003**, *13*, 721–731. [[CrossRef](#)] [[PubMed](#)]
48. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [[CrossRef](#)]

49. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. Dnasp 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)]
50. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2019**, *20*, bbx108. [[CrossRef](#)]
51. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
52. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)]
53. Sharp, P.M.; Li, W.H. The codon Adaptation Index—A measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **1987**, *15*, 1281–1295. [[CrossRef](#)]
54. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)] [[PubMed](#)]
55. Zhang, Z. KaKs_Calculator 3.0: Calculating Selective Pressure on Coding and Non-coding Sequences. *Genom. Proteom. Bioinform.* **2022**, *20*, 536–540. [[CrossRef](#)] [[PubMed](#)]
56. Zhang, Z.; Xiao, J.; Wu, J.; Zhang, H.; Liu, G.; Wang, X.; Dai, L. ParaAT: A parallel tool for constructing multiple protein-coding DNA alignments. *Biochem. Biophys. Res. Commun.* **2012**, *419*, 779–781. [[CrossRef](#)] [[PubMed](#)]
57. Stamatakis, A. Using RAxML to infer phylogenies. *Curr. Protoc. Bioinform.* **2015**, *51*, 6.14.1–6.14.14. [[CrossRef](#)]
58. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.