

## Article

# Characterization of the Plastid Genomes of Four *Caroxylon* Thunb. Species from Kazakhstan

Shyryn Almerikova <sup>1,2</sup>, Moldir Yermagambetova <sup>1</sup> , Bektemir Osmonali <sup>3</sup> , Polina Vesselova <sup>3</sup>, Saule Abugalieva <sup>1,2</sup>  and Yerlan Turuspekov <sup>1,2,\*</sup> 

<sup>1</sup> Institute of Plant Biology and Biotechnology, Almaty 050040, Kazakhstan; almerikovakz@gmail.com (S.A.); ermaganbetova.moldir@bk.ru (M.Y.); absaule@yahoo.com (S.A.)

<sup>2</sup> Faculty of Biology and Biotechnology, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan

<sup>3</sup> Institute of Botany and Phytointroduction, Almaty 050040, Kazakhstan; be96ka\_kz@mail.ru (B.O.); pol\_ves@mail.ru (P.V.)

\* Correspondence: yerlant@yahoo.com

**Abstract:** The family Chenopodiaceae Vent. (Amaranthaceae *s.l.*) is known for its taxonomic complexity, comprising species of significant economic and ecological importance. Despite its significance, the availability of plastid genome data for this family remains limited. This study involved assembling and characterizing the complete plastid genomes of four *Caroxylon* Thunb. species within the tribe Salsoleae *s.l.*, utilizing next-generation sequencing technology. We compared genome features, nucleotide diversity, and repeat sequences and conducted a phylogenetic analysis of ten Salsoleae *s.l.* species. The size of the plastid genome varied among four *Caroxylon* species, ranging from 150,777 bp (*C. nitrarium*) to 151,307 bp (*C. orientale*). Each studied plastid genome encoded 133 genes, including 114 unique genes. This set of genes includes 80 protein-coding genes, 30 tRNA genes, and 4 rRNA genes. Eight divergent regions (*accD*, *atpF*, *matK*, *ndhF-ndhG*, *petB*, *rpl20-rpl22*, *rpoC2*, and *ycf3*) were identified in ten Salsoleae *s.l.* plastid genomes, which could be potential DNA-barcoding markers. Additionally, 1106 repeat elements were detected, consisting of 814 simple sequence repeats, 92 tandem repeats, 88 forward repeats, 111 palindromic repeats, and one reverse repeat. The phylogenetic analysis provided robust support for the relationships within *Caroxylon* species. These data represent a valuable resource for future phylogenetic studies within the genus.

**Keywords:** *Caroxylon*; plastid genome; genome comparison; variable regions; phylogenetic relationships; next-generation sequencing



**Citation:** Almerikova, S.; Yermagambetova, M.; Osmonali, B.; Vesselova, P.; Abugalieva, S.; Turuspekov, Y. Characterization of the Plastid Genomes of Four *Caroxylon* Thunb. Species from Kazakhstan. *Plants* **2024**, *13*, 1332. <https://doi.org/10.3390/plants13101332>

Academic Editors: Khalid Meksem and Yin-Long Qiu

Received: 10 April 2024

Revised: 8 May 2024

Accepted: 9 May 2024

Published: 12 May 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

Chenopodiaceae Vent. (Amaranthaceae *s.l.*) is one of the largest and most ancient plant families inhabiting desert and semi-desert regions worldwide [1,2]. The Chenopodiaceae family encompasses approximately 1700 species distributed among around 110 genera [3]. Ecologically, representatives of the Chenopodiaceae family play crucial roles in desert ecosystems, serving as vital food sources for herbivores and contributing significantly to soil stabilization [4]. One of the family's largest and most important tribes is Salsoleae *s.l.* [5,6]. The precise number of genera in Salsoleae *s.l.* worldwide remains uncertain, with estimates ranging from 32 to 98 [5]. The species of Salsoleae *s.l.* are widely distributed across desert and semi-desert regions spanning Central Asia, the Middle East, Africa, and Europe [7,8]. Among these members is the genus *Caroxylon* Thunb., formerly classified as a section within the *Salsola* L. before being reinstated as a distinct genus [9]. According to POWO (Plants of the World Online) [10], the genus comprises 128 species globally, with nine of them found in Kazakhstan. These species are economically significant as forage plants and sources of medicinal compounds [4,7].

Numerous studies have been conducted to explore the taxonomy of the tribe, including *Caroxylon* species, employing both morphological characteristics and molecular genetics

methodologies [3,6,9,11–18]. However, despite these efforts, a precise taxonomy of the tribe has yet to be established.

Akhani et al. [9] performed an extensive phylogenetic study on the Salsoleae *s.l.* utilizing sequences from the nuclear ribosomal internal transcribed spacer and the plastid *psbB-psbH* region. Certain representatives previously classified under *Salsola s.l.* have been redistributed among reinstated or recently established genera [9]. For instance, *Salsola canescens* (Moq.) Boiss., previously included in the *Salsola* sect. *Belanthera* Iljin under the name *S. boissieri* Botsch. [19] was transferred to the genus *Caroxylon* as *Caroxylon canescens* (Moq.) Akhani [9]. However, Sukhorukov et al. [20] indicated uncertainty regarding the taxonomic transfer of *S. canescens* to the genus *Caroxylon*, and they established a new genus, *Akhania*, for this species based on molecular phylogeny using *ITS* and *psbB-psbH* nucleotide sequences. Furthermore, using three DNA barcodes, Wen and co-authors [6] suggested that all species of tribe Salsoleae *s.l.* were composed of three monophyletic subunits: *Salsola s.str.*, the *Kali* clade, and *Caroxylonea*. Despite the comprehensive phylogenetic study conducted by Akhani et al. [9], not all species from Central Asia, including Kazakhstan, were included in the analysis.

In addition to the numerous phylogenetic analyses conducted on Salsoleae *s.l.* species utilizing nuclear and plastid genome markers, population genetics analysis was also performed. The Amplified Length Polymorphism (AFLP) technique was employed to identify and use AFLP markers to study genetic relationships in four *Salsola* species [21]. Inter simple sequence repeat (ISSR) and start codon targeted (SCoT) molecular markers were used to analyze the genetic relationships between the different species of *Salsola* [22]. Seventeen simple sequence repeat (SSR) markers of *Beta* were employed in the cross-genera amplification of five morphologically distinct invasive *Salsola* taxa [23]. However, only six of them were successfully amplified within the studied *Salsola* taxa [23]. Additionally, several studies were conducted to assess the cross-genera transferability of these SSR markers to *Salsola* species [24,25]. However, there is a notable absence of population structure analyses in representatives of *Caroxylon*.

The SSR markers are valuable genetic markers extensively employed in population studies [26–29]. These markers, composed of tandemly repeated motifs, offer unique advantages in elucidating genetic variation and evolutionary relationships within and among species [30,31]. While SSRs are widespread across the genomes of diverse organisms [32], cpSSRs specifically reside within the chloroplast genome, primarily found in plant cells [26]. Despite the rising popularity of SSR markers in plant population studies, there is a noticeable absence of research on economically significant *Caroxylon* species.

Plastid genome characterization studies of representatives from the tribe Salsoleae *s.l.* have been limited, with only a few published investigations. Specifically, Li et al. [33] examined *Salsola abrotanoides*, while Xie et al. [34] investigated *Caroxylon passerinum*. Comparative analysis of plastid genome data helps reconstruct phylogenetic trees and provides valuable information for understanding the evolutionary relationships among plant species [35,36]. The rapid advancement of next-generation sequencing technology has greatly improved the efficiency and accessibility of obtaining complete plastid genome nucleotide sequences [37]. Sequencing plastid genomes in plants is crucial for advancing various fields of research, including evolutionary biology [38], taxonomy [39], biogeography [40], breeding [41], and conservation [42]. However, despite the widespread use of plastid genome data for comparative analysis, no comparative studies have been conducted on *Caroxylon* plastid genomes.

Using the nomenclature based on molecular evidence [6,9], we studied the sequencing, assembly, and annotation of plastid genomes of four *Caroxylon* species collected in Kazakhstan: *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum*. These species thrive in rocky and clay soils, serving as essential forage for herbivores during autumn and winter. *C. orientale* and *C. nitrarium* are widely distributed throughout Kazakhstan's territory. Furthermore, we conducted a comparative analysis to characterize these plastomes, comparing them with related taxa plastomes available in GenBank. The objectives of this

study encompassed providing plastid genome data for four newly sequenced *Caroxylon* species; comparing the structure of their plastid genomes and identifying variable regions suitable as potential DNA barcoding markers for species identification and phylogenetic analysis; exploring repeat elements, including simple sequence repeats, tandem repeats, forward repeats, palindromic repeats, and reverse repeats within the analyzed plastomes; and utilizing common protein-coding gene sequences for constructing a phylogenetic tree, aiming to clarify the phylogenetic relationships among the studied species.

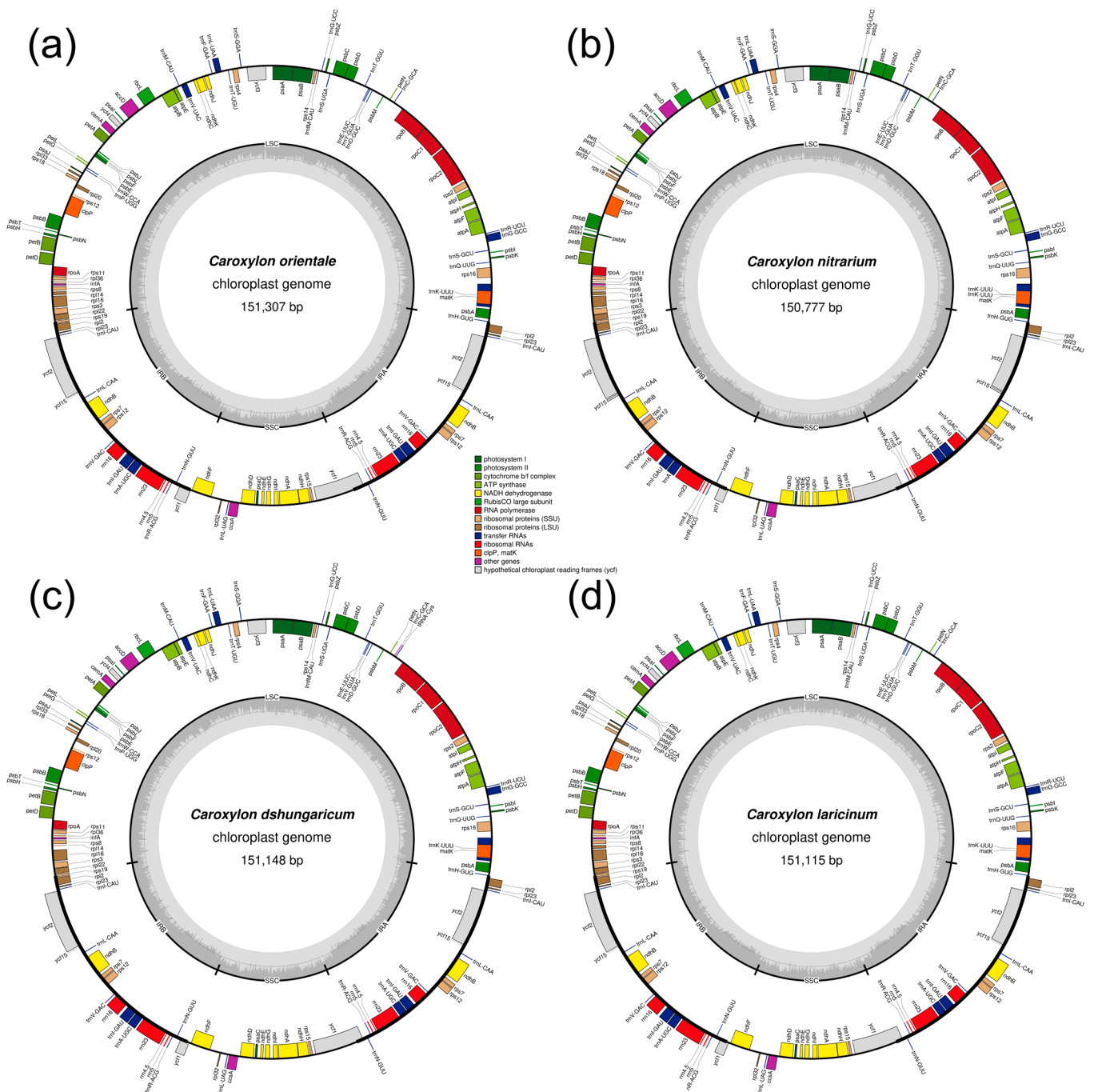
2. Results

2.1. Features of the Plastid Genome

In this study, we sequenced plastid genomes of four *Caroxylon* species (*C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum*) that were collected in Kazakhstan. A total of 27,353,686, 24,456,970, 21,851,266, and 24,687,480 paired-end reads were obtained, each having a sequence length of 151 bp. Subsequently, 25,260,324, 21,693,212, 19,639,670, and 22,024,448 high-quality reads were used for mapping the plastid genome of *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum*, respectively. High-quality data with clean reads totaling over 3.8, 3.5, 3.1, and 3.6 Gb were generated for *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum*, respectively. The sequencing quality values Q20 were determined to be 97.58%, 96.62%, 96.94%, and 96.8% for *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum*, respectively. The Q30 values were 92.11%, 90.07%, 90.76%, and 90.47% for the same species. The newly sequenced plastomes have been submitted to GenBank with the following accession numbers: OR551471 (*C. orientale*), OR552116 (*C. nitrarium*), PP503423 (*C. dzhungaricum*), and PP503424 (*C. laricinum*). The complete plastid genome in four *Caroxylon* species ranged in size from 150,777 bp in *C. nitrarium* to 151,307 bp in *C. orientale* (Table 1). Each of the newly sequenced plastomes exhibited a typical quadripartite structure consisting of four regions: a pair of inverted repeats (IRa and IRb), the large single-copy region (LSC), and the small single-copy region (SSC) (Figure 1). The length of the LSC region varied from 83,329 bp in *C. nitrarium* to 83,706 bp in *C. dzhungaricum*. The SSC region ranged from 18,266 bp in *C. orientale* to 18,999 bp in *C. laricinum*. Additionally, the IR region varied in size from 48,438 bp (*C. laricinum*) to 51,348 bp (*C. orientale*). In terms of GC content, the IR regions displayed the highest GC content, ranging from 42.57% (*C. orientale*) to 42.72% (*C. laricinum*). Following this, the LSC region showed GC content ranging from 34.67% (*C. dzhungaricum*) to 34.69% (*C. laricinum*). Conversely, the SSC region exhibited the lowest GC content, varying from 29.68% (*C. orientale*) to 30.33% (*C. nitrarium*). The overall GC content of the plastome sequences for *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum* was 36.69%, 36.84%, 36.68%, and 36.71%, respectively (Table 1).

**Table 1.** Summary of plastid genome characteristics of *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum*.

	<i>C. orientale</i>	<i>C. nitrarium</i>	<i>C. dzhungaricum</i>	<i>C. laricinum</i>
GenBank numbers	OR551471	OR552116	PP503423	PP503424
Genome size (bp)	151,307	150,777	151,148	151,115
LSC (bp)	83,693	83,329	83,706	83,678
SSC (bp)	18,266	18,986	18,996	18,999
IR (bp)	51,348	48,462	48,446	48,438
Number of total genes	133	133	133	133
Protein-coding genes	80	80	80	80
tRNAs	30	30	30	30
rRNAs	4	4	4	4
Total GC content (%)	36.69	36.84	36.68	36.71
LSC GC content (%)	34.69	34.86	34.67	34.69
SSC GC content (%)	29.68	30.33	28.80	29.82
IR GC content (%)	42.57	42.60	42.70	42.72



**Figure 1.** Plastid genome maps of the four *Caroxylon* species. (a) *C. orientale*, (b) *C. nitrarium*, (c) *C. dzhungaricum*, and (d) *C. laricinum*. Genes from various functional categories are colored accordingly. The darker gray on the inner circle represents GC content, while the lighter gray denotes AT content.

Each of the *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum* plastid genomes encoded 133 genes; 114 of these genes were unique, including 80 protein-coding genes, 30 tRNA genes, and 4 rRNA genes (Table 1). Seven tRNA genes (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*), eight protein-coding genes (*rps7*, *rps12*, *rpl2*, *rpl23*, *ndhB*, *ycf1*, *ycf2*, and *ycf15*), and four rRNA genes (*rrn4.5*, *rrn5*, *rrn16*, and *rrn23*) were identified as duplicated within IR regions. Among the 114 unique genes, 17 contain introns: 6 tRNA genes (*trnA-UGC*, *trnG-GCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, and *trnV-UAC*) and 11 protein-coding genes (*rps12*, *rps16*, *rpl16*, *rpoC1*, *atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *clpP*, and *ycf3*). Notably, *clpP* and *ycf3* stand out as the only genes in this context

with two introns each, while the other 15 genes are characterized by having a single intron each (Table 2).

**Table 2.** Genes identified in plastomes of four *Caroxylon* species.

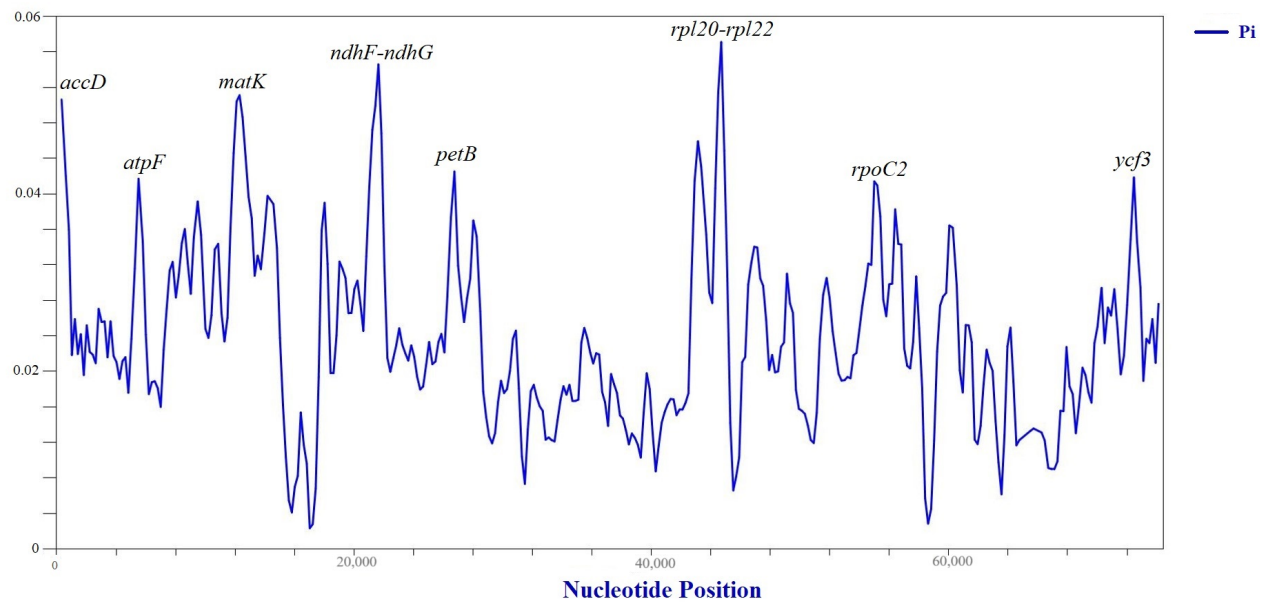
Category	Group of Genes	Name of Genes
Self-replication	Ribosomal RNA	<i>rrn4.5</i> (x2), <i>rrn5</i> (x2), <i>rrn16</i> (x2), <i>rrn23</i> (x2)
	Transfer RNA	<i>trnA</i> -UGC * (x2), <i>trnC</i> -GCA, <i>trnD</i> -GUC, <i>trnE</i> -UUC, <i>trnF</i> -GAA, <i>trnI</i> -CAU, <i>trnG</i> -GCC *, <i>trnG</i> -UCC, <i>trnH</i> -GUG, <i>trnI</i> -CAU (x2), <i>trnI</i> -GAU * (x2), <i>trnK</i> -UUU *, <i>trnL</i> -CAA (x2), <i>trnL</i> -UAA *, <i>trnL</i> -UAG, <i>trnM</i> -CAU, <i>trnN</i> -GUU (x2), <i>trnP</i> -UGG, <i>trnQ</i> -UUG, <i>trnR</i> -ACG (x2), <i>trnR</i> -UCU, <i>trnS</i> -GCU, <i>trnS</i> -GGA, <i>trnS</i> -UGA, <i>trnT</i> -GGU, <i>trnT</i> -UGU, <i>trnV</i> -GAC (x2), <i>trnV</i> -UAC *, <i>trnW</i> -CCA, <i>trnY</i> -GUA
	Small subunit of ribosome	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (x2), <i>rps8</i> , <i>rps11</i> , <i>rps12</i> * (x2), <i>rps14</i> , <i>rps15</i> , <i>rps16</i> *, <i>rps18</i> , <i>rps19</i>
	Large subunit of ribosome	<i>rpl2</i> (x2), <i>rpl14</i> , <i>rpl16</i> *, <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (x2), <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
	RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> *, <i>rpoC2</i>
	Translation initiation factor	<i>infA</i>
Photosynthesis	ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> *, <i>atpH</i> , <i>atpI</i>
	NADH dehydrogenase	<i>ndhA</i> *, <i>ndhB</i> * (x2), <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	Subunits of cytochrome	<i>petA</i> , <i>petB</i> *, <i>petD</i> *, <i>petG</i> , <i>petL</i> , <i>petN</i>
	Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	Photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
	Rubisco	<i>rbcL</i>
Other genes	Maturase	<i>matK</i>
	Protease	<i>clpP</i> **
	Envelope membrane protein	<i>cemA</i>
	Subunit of acetyl-CoA-carboxylase	<i>accD</i>
	C-type cytochrome synthesis gene	<i>ccsA</i>
Genes of unknown function	Hypothetical chloroplast reading frames	<i>ycf1</i> (x2), <i>ycf2</i> (x2), <i>ycf3</i> **, <i>ycf4</i> , <i>ycf15</i> (x2)

\* One-intron-containing genes; \*\* two-intron-containing genes; (x2) duplicated genes.

## 2.2. Plastome Analysis by Sliding Window

We utilized DnaSP 6 software to perform a sliding window analysis aimed at determining the nucleotide diversity (Pi) value within the 80 protein-coding genes in plastid genomes of the *Caroxylon* species. Based on the sequence alignment of common protein-coding genes, we identified eight hypervariable regions: *accD*, *atpF*, *matK*, *ndhF*-*ndhG*, *petB*, *rpl20*-*rpl22*, *rpoC2*, and *ycf3*. Seven of these regions (*accD*, *atpF*, *matK*, *petB*, *rpl20*-*rpl22*, *rpoC2*, and *ycf3*) are situated within the LSC region, while only one region (*ndhF*-*ndhG*) is found in the SSC region (Figure 2). Among the variable regions identified, *rpl20*-*rpl22* exhibited the highest Pi value at 0.05711 (Table 3).

The nonsynonymous (Ka) and synonymous (Ks) substitutions were calculated using DNASP 6 between studied plastid genomes (Table 3). The results suggested that the ratios of Ka/Ks in six out of ten genes listed in Table 3 were >1, indicating that they were under positive selection.



**Figure 2.** Sliding window analysis of the common protein-coding genes (window length: 600 bp; step size: 200 bp) in plastid genomes of the *Caroxylon* species. The vertical axis indicates the nucleotide diversity for each window, while the horizontal axis denotes the midpoint position.

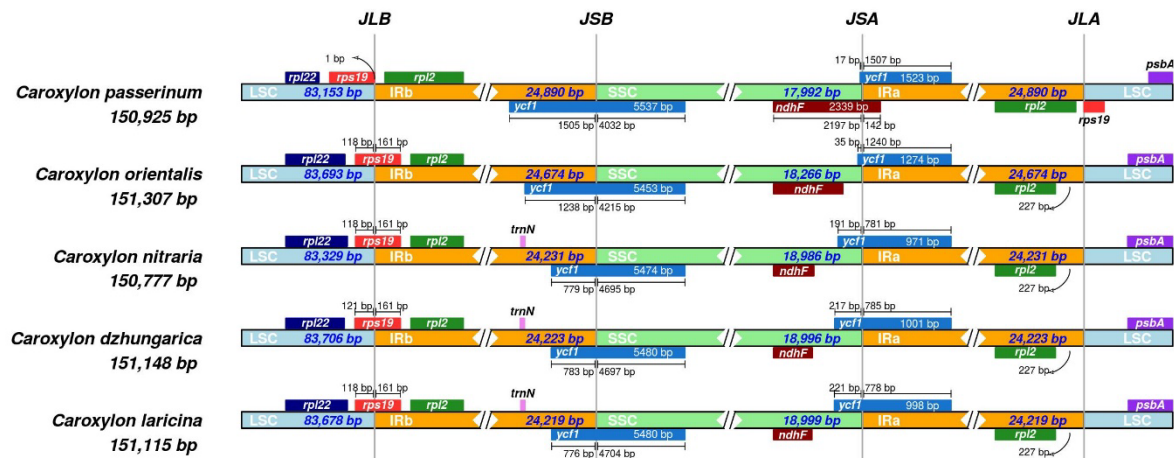
**Table 3.** Regions with highly variable sequences and Ka/Ks ratios of genes in plastid genomes of *Caroxylon* species.

Variable Region	Length	Variable Sites	Parsimony Informative Sites	Nucleotide Diversity	Analysed Genes	Ka/Ks Ratio
<i>accD</i>	723	103	36	0.05063	<i>accD</i>	0.900
<i>atpF</i>	771	75	35	0.04170	<i>atpF</i>	1.308
<i>matK</i>	606	96	37	0.05111	<i>matK</i>	1.813
<i>ndhF-ndhG</i>	606	101	44	0.05459	<i>ndhF</i>	0.600
					<i>ndhG</i>	0.476
<i>petB</i>	721	84	36	0.04252	<i>petB</i>	4.139
<i>rpl20-rpl22</i>	606	114	43	0.05711	<i>rpl20</i>	1.099
					<i>rpl22</i>	2.425
<i>rpoC2</i>	606	78	32	0.04141	<i>rpoC2</i>	0.476
<i>ycf3</i>	672	77	34	0.04185	<i>ycf3</i>	1.104

2.3. IR and SC Regions Boundary Analysis

Our analysis investigated the boundaries of the IR-SSC and IR-LSC regions in the plas-  
tomes of four *Caroxylon* species, comparing them with the reference sample, *C. passerinum*.  
We observed that the gene *rps19* flanked the junction between the LSC and IRb regions in  
the four *Caroxylon* species, whereas in the reference sample *C. passerinum*, the *rps19* gene,  
including a pseudogene, was located within the LSC region. Conversely, the *ndhF* gene in  
*C. passerinum* was positioned to flank the junction between the SSC and IRa regions, while  
in the four *Caroxylon* samples, this gene was situated within the SSC region. The *ycf1* gene  
exhibited a consistent pattern across all five samples, spanning the boundaries of the SSC  
and IRb regions. Additionally, a duplicated copy of the *ycf1* gene was consistently observed  
at the junction between the SSC and IRa regions in each sample (Figure 3).

## Inverted Repeats



**Figure 3.** Comparisons of the borders of LSC, IR, and SSC regions among the five plastid genomes JLB (junctions between LSC and IRb regions), JSB (junctions between IRb–SSC regions), JSA (junctions between SSC–IRa regions), and JLA (junctions between IRa–LSC regions).

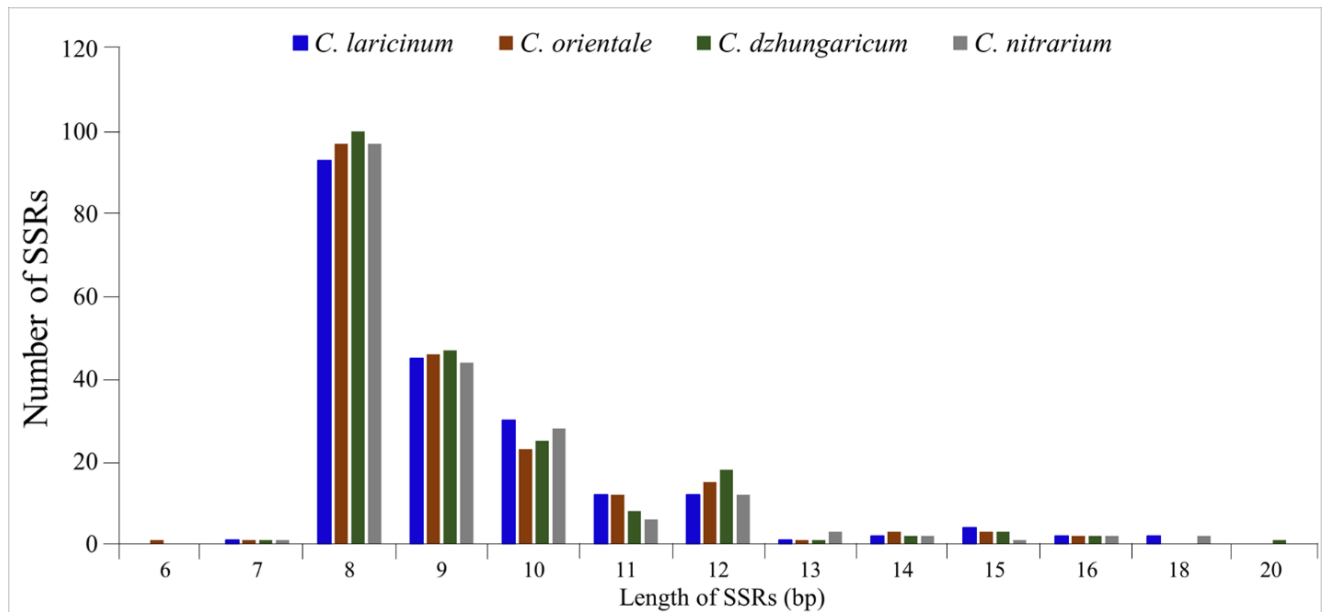
### 2.4. Repeat Sequence Analysis

We identified 814 simple sequence repeats (SSRs) across the four *Caroxylon* plastomes using the MISA tool. SSRs vary among the four *Caroxylon* plastomes, ranging from 198 in *C. nitrarium* to 208 in *C. dzhungaricum*. The analysis revealed that mononucleotide repeats were the most abundant motifs, representing 72.97% of the total SSRs. Dinucleotide repeats followed, accounting for 19.04%, while tetranucleotide repeats constituted 4.18%. Hexanucleotide repeats, representing 0.25% of the total SSRs, were uniquely identified within the *C. loricinum* plastome. The vast majority of mononucleotide repeats consisted of A/T sequences (71.38%), with a minor portion (2.19%) comprising C/G sequences. Regarding dinucleotide repeats, AT/AT repeats were predominant, representing 54.84%, while AC/GT repeats constituted only 6.45%, and AG/CT repeats made up 38.71% (Table 4). Most of the identified SSRs were located in the non-coding and LSC regions of *Caroxylon* plastomes (Supplementary File S1).

**Table 4.** Types and numbers of simple sequence repeats in the plastomes of four *Caroxylon* species.

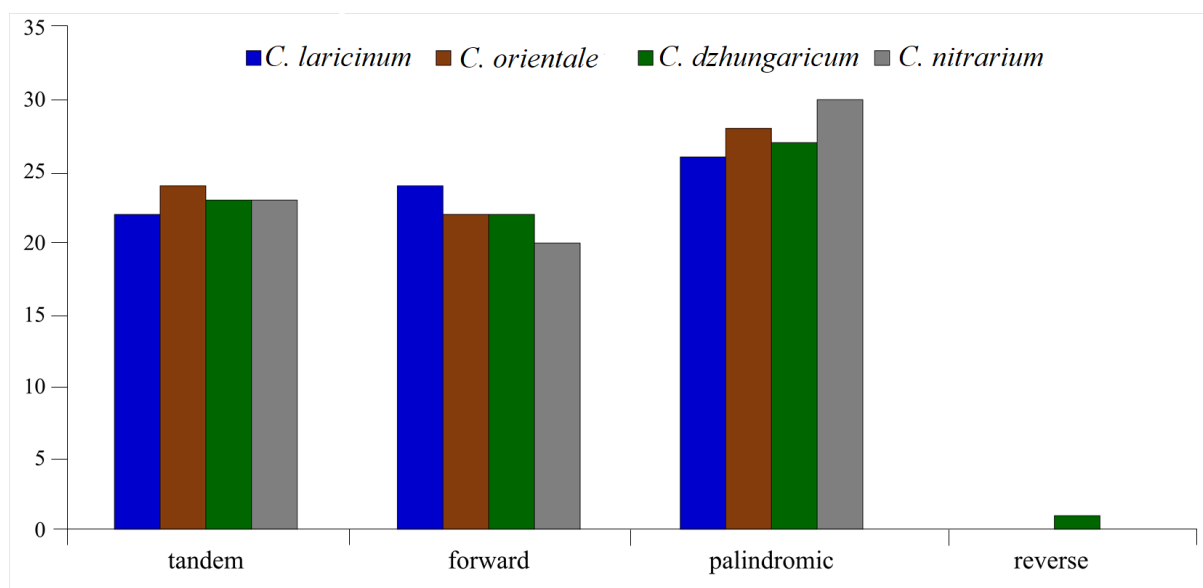
Type	Repeat Unit	<i>C. orientale</i>	<i>C. nitrarium</i>	<i>C.dzhungaricum</i>	<i>C. loricinum</i>	Total (%)	%
Mono-	A/T	148	137	149	147	581 (97.81)	72.97
	C/G	3	4	3	3	13 (2.19)	
	AC/GT	2	4	2	2	10 (6.45)	
Di-	AG/CT	15	15	15	15	60 (38.71)	19.04
	AT/AT	20	25	21	19	85 (54.84)	
	AAG/CTT	-	2	2	2	6 (24)	
Tri-	AAT/ATT	5	4	5	5	19 (76.00)	3.07
	AAAC/GTTT	1	1	1	1	4 (11.76)	
	AAAG/CTTT	3	-	3	2	8 (23.53)	
	AAAT/ATTT	1	1	1	1	4 (11.76)	
Tetra-	AAGG/CCTT	-	1	-	-	1 (2.94)	4.18
	AATC/ATTG	-	-	1	-	1 (2.94)	
	AATT/AATT	2	2	2	2	8 (23.53)	
	ACCT/AGGT	2	2	2	2	8 (23.53)	
Penta-	AAAAG/CTTTT	1	-	1	1	3 (75.00)	0.49
	AAAGG/CCTTT	1	-	-	-	1 (25.00)	
Hexa-	AGCTCC/AGCTGG	-	-	-	2	2 (100.00)	0.25
Total		204	198	208	204	814 (100)	100

The lengths of the identified simple sequence repeats varied from 6 bp in *C. laricinum* to 20 bp in *C. dzhungaricum* plastid genomes. The majority of SSRs were eight base pairs in length, with 93, 97, 100, and 97 repeats identified in the plastid genomes of *C. laricinum*, *C. orientale*, *C. dzhungaricum*, and *C. nitrarium*, respectively. The categorization of repeats with different lengths is presented in Figure 4.



**Figure 4.** The length of the identified simple sequence repeats in plastomes of *C. laricinum*, *C. orientale*, *C. dzhungaricum*, and *C. nitrarium*.

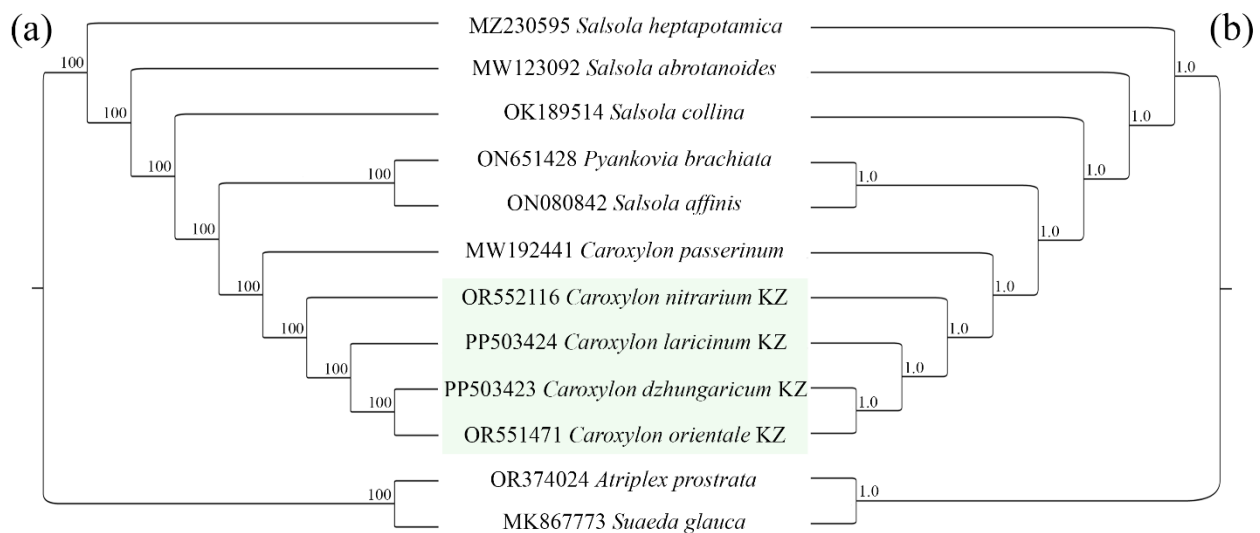
Furthermore, our analysis identified tandem, forward, palindromic, and reverse repeat types in the plastomes of the four *Caroxylon* species. A total of 292 repeats were detected, comprising 92 tandem repeats, 88 forward repeats, 111 palindromic repeats, and just one reverse repeat. The reverse repeat was found solely within the plastome of *C. dzhungaricum* (Figure 5).



**Figure 5.** Number of tandem, forward, palindromic, and reverse repeats in the plastomes of the four *Caroxylon* species.

### 2.5. Phylogenetic Analysis

To elucidate the phylogenetic relationships among the ten species from the Salsoleae tribe, we reconstructed the phylogenetic tree using the Maximum Likelihood (ML) and Bayesian Inference (BI) methods. We utilized nucleotide sequences of 80 protein-coding genes, commonly found in 12 chloroplast genomes, including two outgroup species. Two datasets, (1) nucleotide sequences derived from protein-coding genes and (2) nucleotide sequences encompassing the entire plastid genome, were employed to construct the phylogenetic trees. The phylogenetic analyses based on the ML and BI methods grouped all ten samples into a single clade with strong bootstrap support. The species examined in this study (*C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum*) formed a distinct subclade alongside species obtained from GenBank (*C. passerinum*), thus constituting a subgroup within the *Caroxylon* clade (Figure 6).



**Figure 6.** Phylogenetic tree based on nucleotide sequences from protein-coding genes using the Maximum Likelihood (a) and Bayesian Inference (b) methods. ML bootstrap value and BI posterior probability are given at each node. The species from Kazakhstan were highlighted with the letters “KZ”.

The phylogenetic analyses performed on the entire plastid genome sequences revealed results consistent with those obtained from the dataset of protein-coding gene sequences. The ML and BI trees generated from the complete plastid genome sequences are included in Supplementary File S2.

### 3. Discussion

Data from plastid genomes offer valuable insights for taxonomic studies aimed at assessing evolutionary relationships and conducting comparative analyses across various taxonomic levels [43–46]. In this investigation, we obtained and analyzed four *Caroxylon* plastid genomes using Illumina sequencing technology. The comparative analysis revealed consistent genome structure and gene count across these examined genomes. The study of plastid genomes of *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum* revealed a consistent presence of 133 genes in each species, with 114 genes being unique, including 80 protein-coding genes, 30 tRNA genes, and 4 rRNA genes (Table 1).

Compared with the plastome annotation of *C. passerinum* [34], variations were observed in the number of protein-coding and tRNA genes. Specifically, the total number of protein-coding genes in the *C. passerinum* plastome was reported as 89, including duplicated genes. However, in the four newly sequenced plastomes, this count was 88, as the pseudogene *rps19* was not annotated due to its short length. Conversely, while the *trnG-GCC* and *trnK-UUU* genes were not annotated in the *C. passerinum* plastome [34],

these genes were identified in the four *Caroxylon* species analyzed in this study, resulting in a total of 37 tRNA genes, including duplicated genes.

Our sliding window analysis investigation further identified eight relatively variable regions, encompassing *accD*, *atpF*, *matK*, *ndhF-ndhG*, *petB*, *rpl20-rpl22*, *rpoC2*, and *ycf3* (Figure 2), where six out of ten genes in those regions were under positive selection (Table 3). Noteworthy among these findings is the recognition of the *matK* gene as a core plant barcode by the CBOL Plant Working Group [47]. Moreover, our examination of nucleotide diversity demonstrated similar variability in other species. For instance, Liu et al. [48] and Almerikova et al. [49] identified variability in the *accD* region among *Quercus* and *Juniperus* species, respectively. Rodda and Niissalo [50] observed variability in the *accD* and *ndhF* regions within *Hoya* species plastomes. Additionally, variability in the *rpl20* region among *Aerides* species' plastid genome was noted by Chen et al. [51], while Ding et al. [52] reported variability in the *rpl22* region among the *Clethra* plastid genome. The identified relatively variable regions have the potential to serve as molecular markers for phylogenetic studies of *Caroxylon* species.

The boundaries between IR and SC regions are conserved in studied *Caroxylon* species (Figure 3). The IRb/LSC boundary was consistently found within the *rps19* gene among the four newly sequenced species analyzed in this study. Similarly, the IRb/SSC boundary was located within the *ycf1* gene, while the IRa/SSC boundary was identified within the duplicated copy of the *ycf1* gene. These results are in accordance with previous studies conducted on different species [53,54]. Notably, no major changes in the position of IR regions were observed in the studied plastid genomes, suggesting the absence of substantial structural rearrangements [55,56]. Among more closely related species, any observed shifts in IR boundaries tended to be relatively minor [57].

SSRs are widespread throughout plastid genomes across different species and are extensively employed in plant population studies [58–60]. In this study, the identified SSRs varied in number, with counts ranging from 198 in *C. nitrarium* to 208 in *C. dzhungaricum* plastomes, resulting in a total of 814 SSRs (Table 4). Similar to the previous studies [43,61], mononucleotide repeats were the most prevalent motifs, comprising 594 of the total identified SSRs. Polyadenine (poly-A) or polythymine (poly-T) repeats were the most abundant in the plastid genome of four *Caroxylon* species collected in Kazakhstan, a common phenomenon observed in the plastid genomes of higher plants [62–64]. According to the previous findings [65,66], most of the identified SSRs were located in the non-coding and LSC regions of the four examined plastomes. The SSRs revealed in this analysis offer valuable resources for investigating the population genetics of *Caroxylon* species, thereby filling the gaps in population studies within this genus.

The plastid genome significantly conserves structure and gene composition [57], making it a valuable resource for analyzing phylogenetic relationships across diverse taxonomic levels [67]. Before this study, phylogenetic relationships within the tribe Salsoleae had been assessed using only a limited number of genes, and the precise taxonomy of the tribe remains unresolved. In this analysis, nucleotide sequences from common protein-coding genes of ten representatives of the Salsoleae tribe were employed, comprising four newly sequenced *Caroxylon* species. The phylogenetic tree was reconstructed using the Maximum Likelihood (ML) method. The resulting phylogenetic tree exhibited a topology with high-resolution values at the clades. *Caroxylon* was initially considered a section of the genus *Salsola* [19,68,69], but later it was recognized as a separate genus [9,70]. The monophyly of *Caroxylon* reported in previous studies [6,20] remained consistent with the findings of this study. The assessment of the ML phylogenetic tree suggests that *C. passerinum* (syn. *C. gemmascens*) is the oldest species in the five analyzed taxa. In the four *Caroxylon* species reported in this study, *C. laricinum*, *C. dzhungaricum*, and *C. orientale* have formed a distinct subclade with a high bootstrap value (Figure 6). The fourth species, *C. nitrarium*, seems to have evolved from a common ancestor in the earlier stage of speciation, which is well agreed with the results obtained based on using universal DNA barcodes [6,9].

The robust support values obtained for the phylogenetic relationships inferred from plastid genome data closely mirrored those derived from nuclear gene data, suggesting the reliable resolution of phylogenetic relationships within this genus by plastid genome data. However, further plastid genome data are needed to comprehensively assess the phylogenetic relationships within the *Caroxylon* clade. This study represents the first attempt to evaluate phylogenetic relationships using genomic data in this genus and may serve as a valuable resource for future phylogenetic studies of the genus.

#### 4. Materials and Methods

##### 4.1. Plant Materials and DNA Extraction

Plant leaves of four *Caroxylon* species were collected from Kyzylorda (*C. orientale* and *C. nitrarium*), Almaty (*C. dzhungaricum*), and West Kazakhstan (*C. laricinum*) regions of Kazakhstan (Table 5). The voucher herbarium specimens of *C. orientale*, *C. nitrarium*, *C. dzhungaricum*, and *C. laricinum* were deposited in the Herbarium (AA) of the Institute of Botany and Phytointroduction under accession numbers AA0003263, AA0003264, AA0003265, and AA0003266, respectively. The collected fresh leaves were preserved in silica gel and used for DNA extraction. Total genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) protocol [71], and DNA samples were stored at  $-80^{\circ}\text{C}$  until sequencing.

**Table 5.** Information on collected places of four *Caroxylon* species.

Species	Collected Place	GPS Coordinates
<i>C. orientale</i>	Kyzylorda region, Zhanakorgan district.	44.106111, 67.062778 150 m above sea level (m a. s. l.)
<i>C. nitrarium</i>	Kyzylorda region, Zhanakorgan district.	44.110833, 67.061111 150 m a. s. l.
<i>C. dzhungaricum</i>	Almaty region, Trans-Ili Alatau, left bank of the Charyn river.	43.249722, 78.898611 1210 m a. s. l.
<i>C. laricinum</i>	West Kazakhstan region, Burlinsky district, Bestau village.	51.254444, 53.093611 120 m a. s. l.

##### 4.2. Genome Sequencing, Assembly, and Annotation

The total genomic DNA that passed quality control analysis was used for library preparation using the TruSeq Nano DNA Kit manufactured by Illumina Inc. (San Diego, CA, USA). The plastomes of four *Caroxylon* species were sequenced using an Illumina NovaSeq 6000 platform (Illumina Inc., USA), which was conducted at Macrogen Inc. (Seoul, Republic of Korea). FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>, accessed on 26 January 2024) was used to conduct initial quality control checks on raw sequence data coming from high throughput sequencing pipelines. The Trimmomatic 0.36 software [72] removed adapter sequences from the raw reads. The reads with a quality score over 20 were accepted as good-quality reads. Subsequently, the clean reads were assembled using the NOVOPlasty 4.3.3 program [73]. Further annotation of the assembled sequences was conducted using the published plastome of *Caroxylon passerinum* (MW192441) as a reference. The annotation of protein-coding, rRNA, and tRNA genes was performed using GeSeq [74] with manual corrections. The OrganellarGenomeDRAW 1.3.1 tool (OGDRAW) [75] generated a circular gene map of the four *Caroxylon* species. Finally, the annotated plastome sequences of four *Caroxylon* species were deposited into the GenBank.

##### 4.3. Plastome Analysis by Sliding Window, Ka/Ks Calculation, IR Regions Contraction, and Expansion

To evaluate the nucleotide diversity ( $\Pi$ ) of the plastome sequences, we conducted sliding window analysis using DnaSP 6 software [76]. The window length was configured to 600 bp, with a step size of 200 bp. The synonymous ( $K_a$ ), nonsynonymous ( $K_s$ ), and

Ka/Ks values of protein-coding genes in studied species were analyzed using DnaSP 6 software [76]. The contraction and expansion of the inverted repeat (IR) boundaries in *Caroxylon* species were visualized using IRscope software [77], with *C. passerinum* (MW192441) as the reference.

#### 4.4. Repeat Sequence Analysis

Simple sequence repeats (SSR) were identified in four studied plastome sequences of *Caroxylon* species using the web-based simple sequence repeats finder MISA (<https://webblast.ipk-gatersleben.de/misa/>, accessed on 9 March 2024) tool [78]. The thresholds used were 8 repeat units for mononucleotide, 4 for dinucleotide and trinucleotide, and 3 for tetranucleotide, pentanucleotide, and hexanucleotide SSRs. Analysis of long repeats, including forward (F), reverse (R), and palindromic (P) repeats, was conducted using the REPuter program [79], accessible at <https://bibiserv.cebitec.uni-bielefeld.de/reputer>, accessed on 9 March 2024. This analysis utilized the following parameters: Hamming distance = 3 and a minimum repeat size of 30 base pairs. The Tandem Repeats Finder 4.09 (<https://tandem.bu.edu/trf/trf.html>, accessed on 9 March 2024) tool [80] was utilized to identify tandem repeats (T) with default settings.

#### 4.5. Phylogenetic Analysis

Phylogenetic analysis was conducted using the newly sequenced four *Caroxylon* species, along with six related species and two outgroup species (*Suaeda glauca*, MK867773, and *Atriplex prostrata*, OR374024) obtained from GenBank. Phylogenetic trees were reconstructed using two sets of data: (1) nucleotide sequences from protein-coding genes and (2) nucleotide sequences of the entire plastid genome. Nucleotide sequences from 80 protein-coding genes were utilized, along with outgroups, to construct the phylogenetic tree using the Maximum Likelihood (ML) and Bayesian Inference (BI) methods. The ML phylogenetic tree was constructed using IQ-TREE 2.2.2.6 software [81], employing the best-fit model TVM + F + I + R3, selected based on the Bayesian Information Criterion (BIC). BI analysis was performed using MrBayes [82]. The generated trees were displayed using FigTree 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>, accessed on 10 March 2024).

### 5. Conclusions

The complete plastid genomes of four *Caroxylon* species were sequenced and annotated. The size of the plastid genomes varied from 150,777 in *C. nitrarium* to 151,307 in *C. orientale*. The comparative evaluation of the four plastid genomes indicated that they consisted of 133 genes in each species, including 80 protein-coding genes, 30 tRNA genes, and 4 rRNA genes. The regions *accD*, *atpF*, *matK*, *ndhF-ndhG*, *petB*, *rpl20-rpl22*, *rpoC2*, and *ycf3* were identified as the most divergent regions, with six out of ten genes in those regions being under positive selection. The analysis of four plastid genomes predicted the availability of 814 SSRs, with counts ranging from 198 in *C. nitrarium* to 208 in *C. dzhungaricum*. The ML phylogenetic tree confirmed the monophyletic origin of *Caroxylon*. The assessment of the dendrogram suggested that three *Caroxylon* species (*C. laricinum*, *C. dzhungaricum*, and *C. orientale*) have formed a distinct subclade with a robust genetic relationship. Thus, assessing the complete sequences of four plastid genomes in the genus provided highly informative data for future *Caroxylon* genetic studies.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants13101332/s1>, Supplementary File S1: SSRs list of *Caroxylon* species. Supplementary File S2: Phylogenetic tree based on complete plastid genome data.

**Author Contributions:** Conceptualization, S.A. (Shyryn Almerikova) and Y.T.; Methodology, S.A. (Shyryn Almerikova) and B.O.; Validation, S.A. (Shyryn Almerikova), M.Y. and B.O.; Formal Analysis, S.A. (Shyryn Almerikova) and B.O.; Investigation, S.A. (Shyryn Almerikova), M.Y., B.O. and P.V.; Resources, B.O. and P.V.; Data Curation, S.A. (Shyryn Almerikova) and S.A. (Saula Abugaliyeva); Writing—Original Draft Preparation, S.A. (Shyryn Almerikova) and Y.T.; Writing—Review & Editing,

S.A. (Shyryn Almerikova) and Y.T.; Supervision, S.A. (Shyryn Almerikova) and Y.T.; Project Administration, S.A. (Shyryn Almerikova) and S.A. (Saule Abugalieva); Funding Acquisition, S.A. (Shyryn Almerikova). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP14869593).

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare that they have no competing interests.

## References

- Kadereit, G. Origin and age of Australian Chenopodiaceae. *Org. Divers. Evol.* **2005**, *5*, 59–80. [CrossRef]
- Müller, K.; Borsch, T. Phylogenetics of Amaranthaceae Based on matK/trnK Sequence Data: Evidence from Parsimony, Likelihood, and Bayesian Analyses. *Ann. Mo. Bot. Gard.* **2005**, *92*, 66–102.
- Kadereit, G.; Borsch, T.; Weising, K.; Freitag, H. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C4 photosynthesis. *IJPS* **2003**, *164*, 959–986.
- Osmonali, B.B.; Vesselova, P.V.; Kudabayeva, G.M.; Akhtayeva, N.Z. Phytocenotic features of species of the genus *Salsola* L. (Chenopodiaceae Vent./Amaranthaceae Juss.) in the desert part of the Syrdarya river valley. *Bull. Karaganda Univ. Biol. Med. Geogr. Ser.* **2022**, *1*, 78–85. [CrossRef]
- Kühn, U.; Bittrich, V.; Carolin, R.; Freitag, H.; Hedge, I.C.; Uotila, P.; Wilson, P.G. Chenopodiaceae. In *Flowering Plants·Dicotyledons*; Springer: Berlin/Heidelberg, Germany, 1993; pp. 253–281.
- Wen, Z.B.; Zhang, M.L.; Zhu, G.L.; Sanderson, S.C. Phylogeny of Salsoleae sl (Chenopodiaceae) based on DNA sequence data from ITS, psb B–psb H, and rbc L, with emphasis on taxa of northwestern China. *Plant Syst. Evol.* **2010**, *288*, 25–42. [CrossRef]
- Murshid, S.S.; Atoum, D.; Abou-Hussein, D.R.; Abdallah, H.M.; Hareeri, R.H.; Almukadi, H.; Edrada-Ebel, R. Genus *Salsola*: Chemistry, biological activities and future prospective—A review. *Plants* **2022**, *11*, 714. [CrossRef] [PubMed]
- Abtahi, M.; Zandi Esfahan, E. Effects of phenological stage on forage quality of halophyte species *Salsola arbuscula* Pall. in the central desert of Iran. *Appl. Ecol. Environ. Res.* **2017**, *15*, 901–909. [CrossRef]
- Akhani, H.; Edwards, G.; Roalson, E.H. Diversification of the old world Salsoleae sl (Chenopodiaceae): Molecular phylogenetic analysis of nuclear and chloroplast data sets and a revised classification. *IJPS* **2007**, *168*, 931–956.
- POWO. Plants of the World Online. *Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet. Available online: <http://www.plantsoftheworldonline.org/>* (accessed on 27 April 2024).
- Pyankov, V.I.; Artyusheva, E.G.; Edwards, G.E.; Black, C.C., Jr.; Soltis, P.S. Phylogenetic analysis of tribe Salsoleae (Chenopodiaceae) based on ribosomal ITS sequences: Implications for the evolution of photosynthesis types. *Am. J. Bot.* **2001**, *88*, 1189–1198. [CrossRef]
- Toderich, K.N.; Shuyskaya, E.V.; Ozturk, M.; Juylova, A.; Gismatulina, L.I.L.Y.A. Pollen morphology of some Asiatic species of genus *Salsola* (Chenopodiaceae) and its taxonomic relationships. *Pak. J. Bot.* **2010**, *42*, 155–174.
- Sennikov, A.; Freitag, H. The nomenclatural history of *Salsola sedoides* L. and *Salsola sedoides* Pall. (Suaedoideae and Camphorosmoideae, Chenopodiaceae/Amaranthaceae). *Taxon.* **2014**, *63*, 151–160. [CrossRef]
- Schüssler, C.; Freitag, H.; Koteyeva, N.; Schmidt, D.; Edwards, G.; Voznesenskaya, E.; Kadereit, G. Molecular phylogeny and forms of photosynthesis in tribe Salsoleae (Chenopodiaceae). *J. Exp. Bot.* **2017**, *68*, 207–223. [CrossRef] [PubMed]
- Mosyakin, S.L.; Freitag, H.; Rilke, S. Kali versus *Salsola*: The instructive story of a questionable nomenclatural resurrection. *Isr. J. Plant Sci.* **2017**, *64*, 18–30.
- Feodorova, T.A. Phylogenetic relations of the South African species of *Caroxylon* sect. *Caroxylon* and *Tetragonae* (Chenopodiaceae) based on the morphology and nrITS sequences. *Turczaninowia.* **2011**, *14*, 69–76. (In Russian)
- Feodorova, T.A. Morphological, molecular-phylogenetic taxonomic studies in the genus *Caroxylon* Thunb. sensu latissimo (Caroxyleae, Chenopodiaceae Juss.). In *Leonid Vasil'evich Kudryashov. Ad Memoriam*; Timonin, A.K., Ed.; MAKSPress: Moskva, Russia, 2012; pp. 155–169. (In Russian)
- Mucina, L. *Caroxylon* (Chenopodiaceae s. str.) in continental southern Africa and Madagascar: A preliminary nomenclatural synopsis and biogeographical considerations. *Phytotaxa* **2017**, *312*, 151–178. [CrossRef]
- Botschantzev, V.P. *Collection of Works on Geobotany, Forestry, Paleogeography of Floristics: Two New Genera from the Goosefoot Family*; Academician V. N. Sukachev on the 75th Anniversary; Academy of Sciences of the USSR: Saint Petersburg, Russia, 1956; pp. 108–118. (In Russian)
- Sukhorukov, A.P.; Fedorova, A.V.; Kushunina, M.; Mavrodiev, E.V. Akhania, a new genus for *Salsoladaghestanica*, *Caroxylon-canescens* and *C. carpathum* (Salsoloideae, Chenopodiaceae, Amaranthaceae). *PhytoKeys* **2022**, *211*, 45. [CrossRef] [PubMed]
- Abdel-Hamid, A.M. Characterization of four *Salsola* species and their genetic relationship by AFLP. *Pak. J. Bot.* **2016**, *48*, 1183–1187.
- Alotaibi, M.O.; Abd-Elgawad, M.E. ISSR and ScoT for evaluation of hereditary differences of 29 wild plants in Al Jubail Saudi Arabian. *Saudi J. Biol. Sci.* **2022**, *29*, 3223–3231. [CrossRef] [PubMed]
- McGRAY, H.G.; Ayres, D.R.; Sloop, C.M.; Lee, A.K. Beta SSR loci cross-amplify in five *Salsola* taxa. *Mol. Ecol. Resour.* **2008**, *8*, 608–611. [CrossRef] [PubMed]

24. Welles, S.R.; Ellstrand, N.C. Genetic structure reveals a history of multiple independent origins followed by admixture in the allopolyploid weed *Salsola ryanii*. *Evol. Appl.* **2016**, *9*, 871–878. [\[CrossRef\]](#)
25. Almerikova, S.; Favarisova, N.; Turuspekov, Y.; Abugalieva, S. Cross-Genera Transferability of Microsatellite Markers and Phylogenetic Assessment of Three Species from Western Kazakhstan. *Proc. Latv. Acad. Sci. B Nat. Exact. Appl. Sci.* **2020**, *74*, 325–334. [\[CrossRef\]](#)
26. Ebert, D.; Peakall, R. A new set of universal de novo sequencing primers for extensive coverage of non-coding chloroplast DNA: New opportunities for phylogenetic studies and cpSSR discovery. *Mol. Ecol. Resour.* **2009**, *9*, 777–783. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Meng, F.; Liu, L.; Peng, M.; Wang, Z.; Wang, C.; Zhao, Y. Genetic diversity and population structure analysis in wild strawberry (*Fragaria nubicola* L.) from Motuo in Tibet Plateau based on simple sequence repeats (SSRs). *Biochem. Syst. Ecol.* **2015**, *63*, 113–118. [\[CrossRef\]](#)
28. Wang, M.; Yuan, M.; Li, X.; Wu, X.; Ueno, S.; Cai, M.; Wen, Y. Development and characterization of novel chloroplast microsatellite markers for *Cryptomeria japonica* var. *sinensis* (Cupressaceae) and their cross-species amplification. *Biologia* **2022**, *77*, 361–367. [\[CrossRef\]](#)
29. Yermagambetova, M.; Almerikova, S.; Turginov, O.; Sultangaziev, O.; Abugalieva, S.; Turuspekov, Y. Genetic Diversity and Population Structure of *Juniperus seravschanica* Kom. Collected in Central Asia. *Plants* **2023**, *12*, 2961. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Song, S.-L.; Lim, P.-E.; Phang, S.-M.; Lee, W.-W.; Hong, D.; Prathep, A. Development of chloroplast simple sequence repeats (cpSSRs) for the intraspecific study of *Gracilaria tenuistipitata* (Gracilariales, Rhodophyta) from different populations. *BMC Res. Notes* **2014**, *7*, 77. [\[CrossRef\]](#) [\[PubMed\]](#)
31. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Li, Y.C.; Korol, A.B.; Fahima, T.; Beiles, A.; Nevo, E. Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. *Mol. Ecol.* **2022**, *11*, 2453–2465. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Li, X.; Zhang, Q.; Duo, J.; Yang, Y.; Ju, X.; Duan, R.; Xiong, H. The complete chloroplast genome of *Salsola abrotanoides* (Chenopodiaceae), a desert halophyte shrub in China. *Mitochondrial DNA Part B* **2021**, *6*, 1152–1153. [\[CrossRef\]](#)
34. Xie, W.; Zhang, C.; Wang, Y.; Zhang, Y. The complete chloroplast genome of *Caroxylon passerinum* (Chenopodiaceae), an annual desert plant. *Mitochondrial DNA Part B* **2022**, *7*, 426–427. [\[CrossRef\]](#)
35. Li, L.; Wang, W.; Zhang, G.; Wu, K.; Fang, L.; Li, M.; Zeng, S. Comparative analyses and phylogenetic relationships of thirteen Pholidota species (Orchidaceae) inferred from complete chloroplast genomes. *BMC Plant Biol.* **2023**, *23*, 269. [\[CrossRef\]](#)
36. Xia, Q.; Zhang, H.; Lv, D.; El-Kassaby, Y.A.; Li, W. Insights into phylogenetic relationships in *Pinus* inferred from a comparative analysis of complete chloroplast genomes. *BMC Genom.* **2023**, *24*, 346. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Daniell, H.; Jin, S.; Zhu, X.G.; Gitzendanner, M.A.; Soltis, D.E.; Soltis, P.S. Green giant—A tiny chloroplast genome with mighty power to produce high-value proteins: History and phylogeny. *Plant Biotechnol. J.* **2021**, *19*, 430–447. [\[CrossRef\]](#)
38. Gui, L.; Jiang, S.; Xie, D.; Yu, L.; Huang, Y.; Zhang, Z.; Liu, Y. Analysis of complete chloroplast genomes of *Curcuma* and the contribution to phylogeny and adaptive evolution. *Gene* **2020**, *732*, 144355. [\[CrossRef\]](#)
39. Zhang, Y.; Du, L.; Liu, A.; Hu, W.; Yang, T.J.; Wang, Y. The complete chloroplast genome sequences of five *Epimedium* species: Lights into phylogenetic and taxonomic analyses. *Front. Plant Sci.* **2016**, *7*, 186000. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Ye, J.F.; Niu, Y.T.; Feng, Y.L.; Liu, B.; Hai, L.S.; Wen, J.; Chen, Z.D. Taxonomy and biogeography of *Diapensia* (Diapensiaceae) based on chloroplast genome data. *J. Syst. Evol.* **2020**, *58*, 696–709. [\[CrossRef\]](#)
41. Wu, F.H.; Chan, M.T.; Liao, D.C.; Hsu, C.T.; Lee, Y.W.; Daniell, H.; Lin, C.S. Complete chloroplast genome of *Oncidium Gower Ramsey* and evaluation of molecular markers for identification and breeding in *Oncidiinae*. *BMC Plant Biol.* **2010**, *10*, 68. [\[CrossRef\]](#)
42. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Wu, L.; Nie, L.; Xu, Z.; Li, P.; Wang, Y.; He, C.; Yao, H. Comparative and phylogenetic analysis of the complete chloroplast genomes of three *Paeonia* section *Moutan* species (Paeoniaceae). *Front. Genet.* **2020**, *11*, 558451. [\[CrossRef\]](#)
44. Fan, Y.; Jin, Y.N.; Zhang, K.; Zhou, M. The complete chloroplast genome sequences of eight *Fagopyrum* species: Insights into genome evolution and phylogenetic relationships. *Front. Plant Sci.* **2021**, *12*, 799904. [\[CrossRef\]](#)
45. Wang, C.; Liu, J.; Su, Y.; Li, M.; Xie, X.; Su, J. Complete chloroplast genome sequence of *sonchus brachyotus* helps to elucidate evolutionary relationships with related species of asteraceae. *Biomed Res. Int.* **2021**, *2021*, 9410496. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Cao, Z.; Yang, L.; Xin, Y.; Xu, W.; Li, Q.; Zhang, H.; Xin, P. Comparative and phylogenetic analysis of complete chloroplast genomes from seven *Neocinnamomum* taxa (Lauraceae). *Front. Plant Sci.* **2023**, *14*, 1205051. [\[CrossRef\]](#) [\[PubMed\]](#)
47. CBOL Plant Working Group. A DNA barcode for land plants. *Proc. Nat. Acad. Sci. USA* **2009**, *106*, 12794–12797. [\[CrossRef\]](#)
48. Liu, X.; Chang, E.M.; Liu, J.F.; Huang, Y.N.; Wang, Y.; Yao, N.; Jiang, Z.P. Complete chloroplast genome sequence and phylogenetic analysis of *Quercus bawanglingensis* Huang, Li et Xing, a vulnerable oak tree in China. *Forests* **2019**, *10*, 587. [\[CrossRef\]](#)
49. Almerikova, S.; Yermagambetova, M.; Jumanov, S.; Abugalieva, S.; Turuspekov, Y. Comparative analysis of chloroplast genomes of seven *Juniperus* species from Kazakhstan. *PLoS ONE* **2024**, *19*, e0295550. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Rodda, M.; Niissalo, M.A. Plastome evolution arganizationion in the *Hoya* group (Apocynaceae). *Sci. Rep.* **2021**, *11*, 14520. [\[CrossRef\]](#) [\[PubMed\]](#)

51. Chen, J.; Wang, F.; Zhou, C.; Ahmad, S.; Zhou, Y.; Li, M.; Peng, D. Comparative Phylogenetic Analysis for *Aerides* (Aeridinae, Orchidaceae) Based on Six Complete Plastid Genomes. *Int. J. Mol. Sci.* **2023**, *24*, 12473. [\[CrossRef\]](#)
52. Ding, S.; Dong, X.; Yang, J.; Guo, C.; Cao, B.; Guo, Y.; Hu, G. Complete chloroplast genome of *Clethra fargesii* Franch.; an original sympetalous plant from central China: Comparative analysis, adaptive evolution, and phylogenetic relationships. *Forests* **2021**, *12*, 441. [\[CrossRef\]](#)
53. Du, X.; Zeng, T.; Feng, Q.; Hu, L.; Luo, X.; Weng, Q.; Zhu, B. The Complete Chloroplast Genome Sequence of Yellow Mustard (*Sinapis alba* L.) and Its Phylogenetic Relationship to Other Brassicaceae Species. *Gene* **2020**, *731*, 144340. [\[CrossRef\]](#)
54. Tang, C.; Chen, X.; Deng, Y.; Geng, L.; Ma, J.; Wei, X. Complete chloroplast genomes of *Sorbus sensu stricto* (Rosaceae): Comparative analyses and phylogenetic relationships. *BMC Plant Biol.* **2022**, *22*, 495. [\[CrossRef\]](#)
55. Wicke, S.; Schneeweiss, G.M.; de Pamphilis, C.W.; Muller, 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. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Downie, S.R.; Jansen, R.K. A Comparative Analysis of Whole Plastid Genomes from the Apiales: Expansion and Contraction of the Inverted Repeat, Mitochondrial to Plastid Transfer of DNA, and Identification of Highly Divergent Non-coding Regions. *Syst. Bot.* **2015**, *40*, 336–351. [\[CrossRef\]](#)
57. Zhu, A.; Guo, W.; Gupta, S.; Fan, W.; Mower, J.P. Evolutionary dynamics of the plastid inverted repeat: The effects of expansion, contraction, and loss on substitution rates. *New Phytol.* **2016**, *209*, 1747–1756. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Provan, J.; Powell, W.; Hollingsworth, P.M. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends Ecol. Evol.* **2001**, *16*, 142–147. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Sheller, M.; Ciocirlan, E.; Mikhaylov, P.; Kulakov, S.; Kulakova, N.; Ibe, A.; Curtu, A.L. Chloroplast DNA diversity in populations of *P. sylvestris* L. from Middle Siberia and the Romanian carpathians. *Forests* **2021**, *12*, 1757. [\[CrossRef\]](#)
60. Feng, S.; Jiao, K.; Zhang, Z.; Yang, S.; Gao, Y.; Jin, Y.; Wang, H. Development of chloroplast microsatellite markers and evaluation of genetic diversity and population structure of cutleaf groundcherry (*Physalis angulata* L.) in China. *Plants* **2023**, *12*, 1755. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Cheng, Y.; Zhang, L.; Qi, J.; Zhang, L. Complete chloroplast genome sequence of *Hibiscus cannabinus* and comparative analysis of the Malvaceae family. *Front. Genet.* **2020**, *11*, 227. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Rono, P.C.; Dong, X.; Kirika, P.M.; Hu, G.W.; Wang, Q.F. Initial complete chloroplast genomes of *Alchemilla* (Rosaceae): Comparative analysis and phylogenetic relationships. *Front. Genet.* **2020**, *11*, 560368. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Liang, D.; Wang, H.; Zhang, J.; Zhao, Y.; Wu, F. Complete chloroplast genome sequence of *Fagus longipetiolata* Seemen (Fagaceae): Genome structure, adaptive evolution, and phylogenetic relationships. *Life* **2022**, *12*, 92. [\[CrossRef\]](#)
64. Yermagambetova, M.; Abugalieva, S.; Turuspekova, Y.; Almerikova, S. Illumina sequencing data of the complete chloroplast genome of rare species *Juniperus seravschanica* (Cupressaceae) from Kazakhstan. *Data Brief* **2023**, *46*, 108866. [\[CrossRef\]](#)
65. Wang, W.; Yang, T.; Wang, H.L.; Li, Z.J.; Ni, J.W.; Su, S.; Xu, X.Q. Comparative and phylogenetic analyses of the complete chloroplast genomes of six almond species (*Prunus* spp. L.). *Sci. Rep.* **2020**, *10*, 10137. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Yanfei, N.; Tai, S.; Chunhua, W.; Jia, D.; Fazhong, Y. Complete chloroplast genome sequences of the medicinal plant *Aconitum transsectum* (Ranunculaceae): Comparative analysis and phylogenetic relationships. *BMC Genom.* **2023**, *24*, 90. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Daniell, H.; Lin, C.S.; Yu, M.; Chang, W.J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Ulbrich, E. *Chenopodiaceae. Die Natürlichen Pflanzenfamilien*; Engelmann: Leipzig, Germany, 1934; Volume 16, pp. 379–584.
69. Freitag, H. *Salsola* L. (*Chenopodiaceae*); with a contribution from S. Rilke; Akademische Druck und Verlagsanstalt: Graz, Austria, 1997; Volume 172, pp. 154–255.
70. Tzvelev, N.N. Notes on Chenopodiaceae of Eastern Europe. *Ukr. Bot. Zhurnal* **1993**, *50*, 78–85. (In Russian)
71. Doyle, J.J.; Doyle, J.L. A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
72. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Dierckxsens, N.; Mardulyn, P.; NOVOPlasty, G.S. De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **2017**, *45*, e18. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Lohse, M.; Drechsel, O.; Bock, R. OrganellarGenomeDRAW (OGDRAW): A tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* **2007**, *52*, 267–274. [\[CrossRef\]](#)
76. 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. [\[CrossRef\]](#)
77. Amiryousefi, A.; Hyvönen, J.; Pocza, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Beier, S.; Thiel, T.; Munch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [\[CrossRef\]](#) [\[PubMed\]](#)

79. 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](#)] [[PubMed](#)]
80. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [[CrossRef](#)] [[PubMed](#)]
81. Nguyen, L.T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)]
82. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Syst. Biol.* **2012**, *61*, 539–542. [[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.