

Article

Chloroplast Genome of *Salvia* Sect. *Drymosphace*: Comparative and Phylogenetic Analysis

Ting Su ^{1,2}, Yan-Fei Geng ³, Chun-Lei Xiang ⁴ , Fei Zhao ⁴, Mei Wang ^{1,2}, Li Gu ^{1,2} and Guo-Xiong Hu ^{1,2,*}

¹ College of Life Sciences, Guizhou University, Guiyang 550025, China; sutingyo@163.com (T.S.); wmgzu20@163.com (M.W.); guligz@163.com (L.G.)

² The Key Laboratory of Plant Resources Conservation and Germplasm Innovation in Mountainous Region Ministry of Education, Guizhou University, Guiyang 550025, China

³ College of Tea Science, Guizhou University, Guiyang 550025, China; yfgeng@gzu.edu.cn

⁴ CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China; xiangchunlei@mail.kib.ac.cn (C.-L.X.); zhaofei@mail.kib.ac.cn (F.Z.)

* Correspondence: gxhu@gzu.edu.cn

Abstract: Sect. *Drymosphace* is one of eight sections of *Salvia* subg. *Glutinaria* and includes 13 species and one dubious species that hold great economic value. Although the section is well supported, interspecific relationships remain unresolved. Moreover, most of this section's plastome information remains unknown. In this study, we sequenced and assembled eight sect. *Drymosphace* plastomes and conducted comparative analyses within this section. The length of plastid genome sequences ranged from 151,330 bp to 151,614 bp, with 80 protein-coding, 30 tRNA, and four rRNA genes being annotated. The plastomes were found to be as conservative as other Lamiaceae species, showing high consistency and similarity in terms of gene content, order, and structure. Within the sect. *Drymosphace*, single-copy regions were more variable than IR regions, and the intergenic regions were more variable than the coding regions; nine hypervariable regions were detected, and some of them may be useful for the phylogenetic analysis of *Salvia*. The topologies inferred from all of the data sets indicated that sect. *Drymosphace* was monophyletic and that *S. honania* was sister to *S. meiliensis*. Compared to previous studies involving more sect. *Drymosphace* species, phylogenomic analyses can improve the phylogenetic resolution considerably.

Keywords: *Salvia miltiorrhiza*; subg. *Glutinaria*; plastid genome; phylogenomics



Citation: Su, T.; Geng, Y.-F.; Xiang, C.-L.; Zhao, F.; Wang, M.; Gu, L.; Hu, G.-X. Chloroplast Genome of *Salvia* Sect. *Drymosphace*: Comparative and Phylogenetic Analysis. *Diversity* **2022**, *14*, 324. <https://doi.org/10.3390/d14050324>

Academic Editors: Hong-Hu Meng, Yi-Gang Song and Mario A. Pagnotta

Received: 30 January 2022

Accepted: 15 April 2022

Published: 22 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 non-monophyletic nature of traditionally defined *Salvia* led to the establishment of a broad definition of *Salvia* that reduces the five small embedded genera (*Rosmarinus*, *Perovskia*, *Dorystaechas*, *Meriandra* and *Zhumeria*) to subgenera, resulting in a total of 11 subgenera being recognized within *Salvia* [1–3]. *Salvia* is the largest genus in Lamiaceae and includes approximately 1000 species. The genus has a subcosmopolitan distribution but is mainly concentrated in South America, Southwest Asia and the Mediterranean region, and East Asia [3]. As one of biodiversity centers of *Salvia*, approximately 100 species have been recorded in East Asia, 85 of which are native to China [4–11]. Based on recent studies, with the exception of *Salvia grandifolia* and *S. deserta* belonging to subg. *Sclarea*, the rest of East Asian *Salvia* have been placed into the newly established subg. *Glutinaria* [2,12].

Sect. *Drymosphace* was first established by George Bentham in 1832–1836 [13] and was later placed in subg. *Salvia* in 1876 [14]. Based on stamen morphology, Stibal [15] transferred sect. *Drymosphace* to subg. *Sclarea*. Following Stibal, Wu [16] also classified sect. *Drymosphace* into subg. *Sclarea* and established three series (ser. *Miltiorrhizae*, ser. *Plectranthoidites*, and ser. *Honaniae*) within this section. According to Wu's circumscription, a total of 19 species and three varieties can be recognized in sect. *Drymosphace* [4,5,9,16],

and all of these taxa are endemic to China, with the exception of *S. nubicola* and *S. plectranthoides*, which extend to some Himalayan countries. Recent phylogenetic studies have demonstrated that sect. *Drymosphace* and the three series sensu Wu [16] are not monophyletic [2,3]. Integrating morphological with molecular phylogenetic evidence, the redefined sect. *Drymosphace* was transferred from subg. *Sclarea* to the newly established subg. *Glutinaria* due to its synapomorphy of robust taproots, pinnate leaves, relatively long corollas (length > 2 cm), and fused deformed posterior thecae [2]. According to the synapomorphy, five species were removed from the previous sect. *Drymosphace* sensu Wu of subg. *Glutinaria* [16]: *S. cavaleriei* and *S. prionitis* were transferred to the newly established sect. *Sobiso*, *S. petrophila* was transferred to sect. *Sonchifoliae*, *S. nubicola* was transferred to sect. *Glutinaria*, *S. trijuga* was transferred to sect. *Substoloniferae*, and *S. brevicconnectivata* was regarded as a dubious species [2].

Sect. *Drymosphace* are importance of medical value, with *Salvia miltiorrhiza* being most widely used. “Danshen”, a famous traditional Chinese medicine originating from the dried roots of *S. miltiorrhiza*, has been extensively applied to treat coronary heart diseases and cerebrovascular diseases [17]. Due to mass market demand, the wild resource of *S. miltiorrhiza* has reduced sharply. However, although sect. *Drymosphace* is well supported by molecular and morphological evidence, interspecific relationships remain unresolved. Moreover, *S. miltiorrhiza* and its closely related species are morphologically similar, which makes it difficult to identify authentic products, and an increasing number of substitutes and adulterants have been found on the market.

Currently, phylogenetic studies based on short DNA fragments have only been able to determine the phylogenetic backbone of *Salvia* [2,18–22], and relationships among species require further focus. The shallow rate of DNA evolution is one of the main obstacles to determining the relationships between low-level taxa. Due to there being a limited amount of information, the resolutions of phylogenetic trees that has been inferred from chloroplast gene fragments are often low [2,23]. With the birth of next-generation sequencing (NGS) technology, genomics has entered an era of low-cost, large-scale, and high-throughput sequencing, and an increasing number of genomic data are being used for phylogenetic analysis. As one of the three major genomes of green plants, the genomic structure and gene content of chloroplasts is relatively conservative and easy to sequence. At present, plastid genome data are being used to solve phylogenetic relationships at different taxonomic levels and have been demonstrated to be effective [24–26]. Although attempts to utilize NGS to analyze plastome characterization and even though the phylogeny of *Salvia* have been determined, few taxa of sect. *Drymosphace* were involved in these studies [27,28].

In this study, we sequenced eight plastomes of sect. *Drymosphace* and conducted comparative genomics analyses together with two other data sets that have been published on this section (*Salvia meiliensis* and *S. miltiorrhiza*). The aims of this study were (1) to characterize the plastome structure within sect. *Drymosphace*; (2) to screen hyper-variable regions of sect. *Drymosphace*; and (3) to infer the phylogeny of sect. *Drymosphace*.

2. Materials and Methods

2.1. Plant Materials, DNA Extraction and Genome Sequencing

In this study, 14 taxa (13 species and one variety) were sampled from *Salvia* subg. *Glutinaria*, including 10 taxa from sect. *Drymosphace* (Table 1). Eight sect. *Drymosphace* plastomes were newly sequenced, and the other sequences were downloaded from the GenBank database. Additionally, three species from other subgenera were selected as outgroups (Table 1).

The total genomic DNA was extracted from fresh or silica gel-dried leaves using the modified CTAB method [29]. In order to obtain qualified DNA for library construction, a Qubit Fluorometer was used to determine and calculate the DNA concentration and yield when they were at least $c > 12.5$ ng/ μ L and $m > 1$ μ g, and the sample integrity was assessed by means of electrophoresis on a 1% agarose gel. Illumina Hiseq-2500 platform at

BGI-Wuhan was utilized for sequencing by paired-end (PE) library of 150 bp, and 3 to 5 GB of raw short sequence data were generated for each species.

Table 1. Voucher information and GenBank and SRA accession numbers for taxa used in this study.

Taxa	Subgenus	Voucher	Locality	GenBank Accession	SRA Accession
<i>S. bowleyana</i>	<i>Glutinaria</i>	GX Hu and F Zhao 131	China, Fujian	MW435404 *	SRR18650098 *
<i>S. bulleyana</i>	<i>Glutinaria</i>	NA	NA	MH603954	NA
<i>S. chanryoenica</i>	<i>Glutinaria</i>	s.n. (KH)	South Korea	MH261357	NA
<i>S. dabieshanensis</i>	<i>Glutinaria</i>	GX Hu and F Zhao 0165	China, Anhui	MW435405 *	SRR18587923 *
<i>S. honania</i>	<i>Glutinaria</i>	GX Hu and F Zhao 0168	China, Henan	MW435406 *	SRR18600490 *
<i>S. meiliensis</i>	<i>Glutinaria</i>	GX Hu and FZ Shangguan Hu0089	China, Anhui	MN520018	NA
<i>S. miltiorrhiza</i>	<i>Glutinaria</i>	Cultivated	China, Beijing	HF586694	NA
<i>S. nanchuanensis</i>	<i>Glutinaria</i>	JX Yang and XZ He YJX-01	China, Hunan	MW435407 *	SRR18595850 *
<i>S. nanchuanensis</i> <i>var. pteridifolia</i>	<i>Glutinaria</i>	GX Hu and B Pan 615	China, Guangxi	MW435408 *	SRR18595337 *
<i>S. plecranthoides</i>	<i>Glutinaria</i>	GX Hu et al. 0006	China, Yunnan	MW435409 *	SRR18650097 *
<i>S. prattii</i>	<i>Glutinaria</i>	NA	China, Yunnan	MK944407	NA
<i>S. przewalskii</i>	<i>Glutinaria</i>	NA	NA	MH603953	NA
<i>S. subbipinnata</i>	<i>Glutinaria</i>	JX Yang and XZ He YJX-04	China, Zhejiang	MW435410 *	SRR18650096 *
<i>S. yunnanensis</i>	<i>Glutinaria</i>	GX Hu 603	China, Guizhou	MW435411 *	SRR18595336 *
<i>S. hispanica</i>	<i>Calosphace</i>	Cultivated, SD118	China, Shandong	MT083896	NA
<i>S. rosmarinus</i>	<i>Rosmarinus</i>	Cultivated	China, Shaanxi	KR232566	NA
<i>S. officinalis</i>	<i>Salvia</i>	So_003	NA	MG772529	NA

NA = information is unavailable. * = newly sequenced plastomes in this study. SRA = Sequence Read Archive.

2.2. Plastome Assembly and Annotation

After filtering the raw data and discarding the low-quality reads, the remaining PE reads were assembled into whole plastomes using GetOrganelle v1.6.2a [30]. The assembly graph of the generated complete plastome was verified by Bandage v.0.8.1 [31]. Plastome sequences were annotated using software PGA [32], using the *Salvia miltiorrhiza* plastome [33] as reference. The tRNA genes were further verified using the online tRNAscan-SE [34] search servers and then manually adjusted in Geneious 10.0.5 [35]. OrganellarGenomeDRAW, an online program, was used to generate circular annotated plastid genome maps [36], and the plastomes were deposited in the GenBank database (Table 1).

2.3. Codon Usage and Repeated Sequence Analysis

Relative Synonymous Codon Usage (RSCU) is a parameter that is used to evaluate the codon usage preferences of protein-coding sequences. Here, the RSCU values for all of the protein-coding sequences were computed using the program codon W 1.4.4 [37]. The codon usage in the form of heatmap was conducted by employing R language with the RSCU value. An RSCU value > 1 indicates that codon usage is highly preferred, an RSCU value = 1 means that codon usage is not preferred, and an RSCU value < 1 indicates that the codon usage is low [38].

The simple sequence repeat (SSR) of sect. *Drymosphace* plastomes were identified using MISA [39] by setting the minimum number of repeat units to 8, 4, 4, 3, 3, and 3 for the mono-, di-, tri-, tetra-, penta-, and hexanucleotides, respectively. Four types of dispersed repeats (forward, reverse, palindrome, and complement repeats) in sect. *Drymosphace* were

determined using REPuter [40] with a repeat size ≥ 30 bp and sequence identities ≥ 90 (Hamming distance of 3).

2.4. Comparative Genomic Analyses

The boundaries of single-copy (SC) and reverse repeat (IR) regions of sect. *Drymosphace* were determined in Geneious 10.0.5, and the boundary expansion and contraction sketch of IR was drawn using Adobe Illustrator CC. mVISTA [41] was used to determine the variability in the complete plastome sequences of sect. *Drymosphace* in Shuffle-LAGAN mode, and *Salvia bowleyana* was selected as the reference genome. Genome rearrangement was carried out by Mauve [42] using *S. bowleyana* as a reference genome. DnaSP v.5.0 [43] was employed to analyze nucleotide diversity (Pi) within sect. *Drymosphace* by setting the step size to 200 bp with a 600 bp window length.

2.5. Phylogenetic Analysis

To infer the phylogenetic relationships of sect. *Drymosphace*, 14 plastid genomes from subg. *Glutinaria* were selected to carry out analyses with *S. rosmarinus*, *S. hispanica*, and *S. officinalis* from three other subgenera as outgroups. Seven matrices, including complete plastome (CP) sequences with one IR region excluded, large single copy (LSC), small single copy (SSC), IR, protein-coding exons (coding regions, CR), intergenic spacer and introns (non-coding region, NCR), and nine hyper-variable region (HVR) data sets were prepared for the phylogenetic analyses. The sequences were aligned with MAFFT [44] using Gblocks v0.91 [45] to exclude the ambiguously aligned positions. The neat sequence matrices were employed to infer phylogenetic trees using both maximum likelihood (ML) and Bayesian inference (BI).

Under the GTRGAMMA substitution model, ML analyses were performed using RAxML-HPC2 on XSEDE v.8.2.12 [46] on the CIPRES Science Gateway (<http://www.phylo.org/> (accessed on 20 March 2022)) [47]. With the exception of setting bootstrap iterations ($-\#I-N$) to 1000, the other parameters used default values.

BI analysis was carried out in MrBayes 3.2.6 [48] and implemented in PhyloSuite [49]. The ModelFinder [50] was utilized to select the best model according to the Akaike information criterion (AIC) (Table 2). Four Markov chain Monte Carlo (MCMC) iterations were run simultaneously for 2,000,000 generations. Each run started with a random tree, and a random tree was sampled every 1000 generations. Stationarity was considered to be reached when the average standard deviation of the split frequencies was less than 0.01. After discarding the first 25% trees as burn-in, the remaining trees were used to calculate a majority-rule consensus tree for each matrix.

Table 2. Summary of the data set information for phylogenetic analyses.

Data Matrix	Aligned Length (bp)	Constant Sites (bp)	Variable Sites (bp)	Parsimony Informative (bp)	Best Fit Model (AIC)
CP	151,218	144,550 (95.59%)	6668	2366 (1.56%)	GTR + F + I + G4
LSC	82,556	77,911 (94.37%)	4645	1637 (1.98%)	GTR + F + I + G4
SSC	17,531	15,967 (91.08%)	1564	567 (3.23%)	GTR + F + G4
IR	25,563	25,335 (99.11%)	228	80 (0.31%)	GTR + F + I
CR	66,888	64,304 (96.14%)	2584	921 (1.38%)	GTR + F + I + G4
NCR	59,877	56,165 (93.80%)	3712	1346 (2.25%)	GTR + F + I + G4
HVR	10,298	9096 (88.33%)	1064	363 (3.52%)	GTR + F + G4

3. Results and Discussion

3.1. Plastome Features of Sect. *Drymosphace*

The length of complete plastomes of sect. *Drymosphace* ranged from 151,330 bp (*Salvia dabieshanensis*) to 151,614 bp (*S. meiliensis*). All of the plastomes displayed a typical quadripartite architecture that contained a large single copy (LSC: 82,629–82,854 bp) and a small

single copy (SSC: 17,541–17,587 bp) separated by two copies of an inverted repeat (IR: 25,520–25,590 bp). The total GC content varied slightly from 38.00% to 38.03%, and the IR regions had the highest GC content (43.10–43.14%) followed by those in the LSC (36.12–36.15%) and SSC (31.97–32.05%) regions (Figure 1, Table 3).

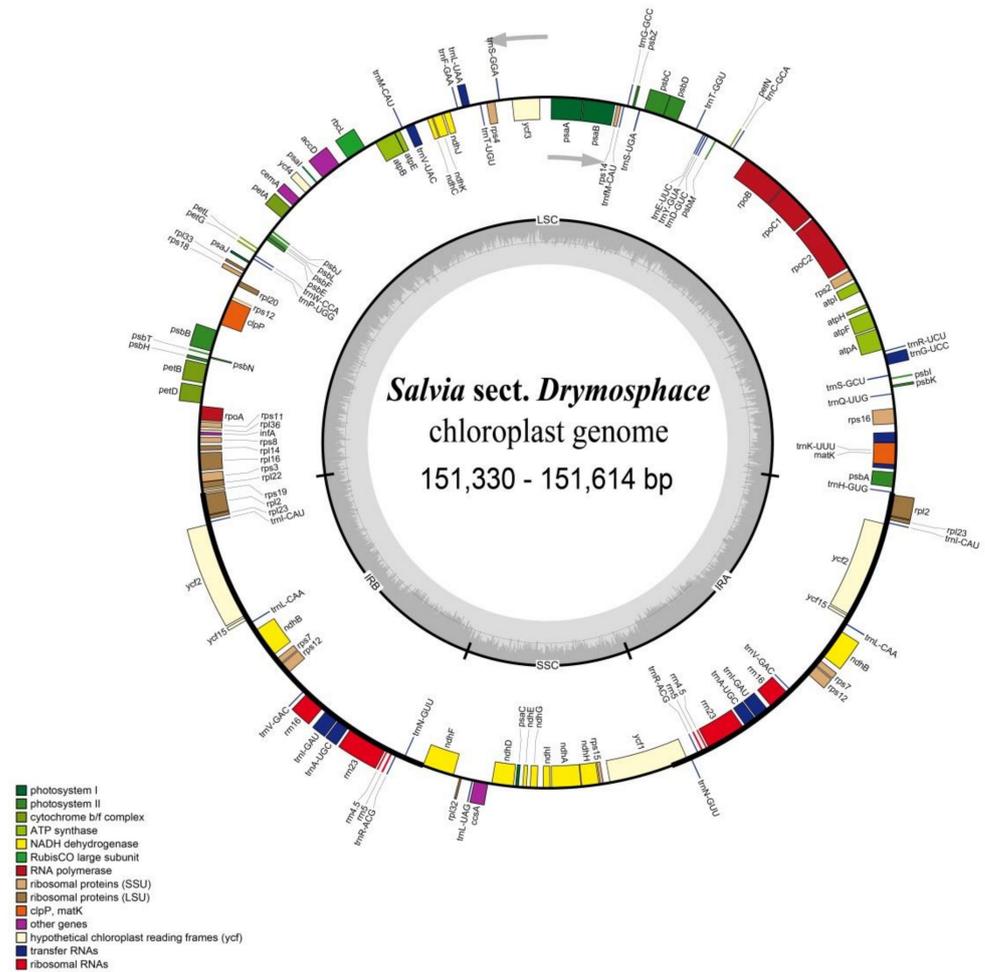


Figure 1. Complete plastome gene map of *Salvia sect. Drymosphace*. Genes outside of the circle are transcribed in the counterclockwise direction, and those that are inside are transcribed in the clockwise direction. LSC—large single copy; SSC—small single copy; IR—inverted repeat.

A total of 114 unique genes, including 80 protein-coding, 30 tRNA, and four rRNA genes, were detected in each species (Table 3). Four rRNA, seven tRNA, and seven protein-coding genes were duplicated in IR regions (Table 3). Of the 18 genes containing intron, 15 comprised only one intron, and three genes harbored two introns. The *rps12* gene was recognized as being a trans-spliced gene, with the 5' end located in the LSC region and the duplicated 3' end in the IR region (Figure 1, Table 4).

Table 3. Plastome features of 10 taxa of *Salvia* sect. *Drymosphace* presented in this study.

Taxa	Clean Reads	Reads Used for Assembly	Mean Coverage (×)	Complete		LSC		SSC		IR		Number of Genes	PCG	tRNA Genes	rRNA Genes
				Size(bp)	GC (%)	Length (bp)	GC (%)	Length (bp)	GC (%)	Length (bp)	GC (%)				
<i>S. bowleyana</i>	6961,322	6,160,014	526	151,508	38.01	82,809	36.14	17,585	32.01	25,557	43.12	114	80	30	4
<i>S. dabieshanensis</i>	11,727,732	9,644,622	504	151,330	38.02	82,629	36.14	17,587	32.01	25,557	43.12	114	80	30	4
<i>S. honania</i>	7,221,162	6,369,236	514	151,522	38.01	82,768	36.13	17,586	31.97	25,584	43.11	114	80	30	4
<i>S. meiliensis</i>	68,112,046	25,577,470	1970	151,614	38.00	82,854	36.12	17,580	32.00	25,590	43.10	114	80	30	4
<i>S. miltiorrhiza</i>	NA	NA	NA	151,332	38.02	82,698	36.15	17,556	32.01	25,539	43.12	114	80	30	4
<i>S. nanchuanensis</i>	9,188,568	7,579,704	507	151,426	38.02	82,791	36.14	17,563	31.99	25,536	43.13	114	80	30	4
<i>S. nanchuanensis</i> <i>var. pteridifolia</i>	15,948,806	13,152,169	514	151,455	38.01	82,831	36.12	17,584	32.05	25,520	43.14	114	80	30	4
<i>S. plectranthoides</i>	6,978,570	6,231,469	512	151,416	38.01	82,775	36.13	17,563	32.03	25,539	43.12	114	80	30	4
<i>S. subbipinnata</i>	5,368,560	4,649,062	516	151,388	38.03	82,769	36.15	17,541	32.02	25,539	43.13	114	80	30	4
<i>S. yunnanensis</i>	6,448,098	5,394,701	508	151,413	38.02	82,652	36.14	17,581	32.03	25,590	43.11	114	80	30	4

Table 4. Genes present in the plastomes of 10 taxa of sect. *Drymosphace*.

Gene Functions	Group of Genes	Name of Genes
Photosynthesis	Subunits of ATP synthase	<i>atpA, atpB, atpE, atpF *</i> , <i>atpH, atpI</i>
	Subunits of NADH dehydrogenase	<i>ndhA *</i> , <i>ndhB *</i> (×2), <i>ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	Subunits of cytochrome	<i>petA, petB *</i> , <i>petD *</i> , <i>petG, petL, petN</i>
	Subunits of photosystem I	<i>psaA, psaB, psaC, psaI, psaJ</i>
	Subunits of photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
	Subunit of rubisco	<i>rbcL</i>
Self-replication	Large subunit of ribosome	<i>rpl2 *</i> (×2), <i>rpl14, rpl16 *</i> , <i>rpl20, rpl22, rpl23</i> (×2), <i>rpl32, rpl33, rpl36</i>
	DNA-dependent RNA polymerase	<i>rpoA, rpoB, rpoC1 *</i> , <i>rpoC2</i>
	Small subunit of ribosome	<i>rps2, rps3, rps4, rps7</i> (×2), <i>rps8, rps11, rps12 **</i> (×2), <i>rps14, rps15, rps16*, rps18, rps19</i>
	rRNA Genes	<i>rrn4.5</i> (×2), <i>rrn5</i> (×2), <i>rrn16</i> (×2), <i>rrn23</i> (×2)
	tRNA Genes	<i>trnA-UGC *</i> (×2), <i>trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-M-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>
Other genes	Subunit of Acetyl-CoA-carboxylase	<i>accD</i>
	c-type cytochrome synthesis gene	<i>ccsA</i>
	Envelop membrane protein	<i>cemA</i>
	Protease	<i>clpP **</i>
	Translational initiation	<i>infA</i>
	Maturase	<i>matK</i>
Unknown function	Conserved open reading frames	<i>ycf1, ycf2</i> (×2), <i>ycf3 **</i> , <i>ycf4, ycf15</i> (×2)

* gene with a single intron; ** gene with two introns; (×2) duplicated gene.

The plastid genomes of sect. *Drymosphace* showed high consistency and similarity in terms of their gene structure, order, and content. These results are consistent with other plastomes of Lamiaceae [26–28,33,51] and are also as conservative as most of the angiosperms that have been previously reported [52–56].

3.2. Amino Acid Abundance and Codon Usage

The number of codons in sect. *Drymosphace* ranged from 26,501 (*Salvia miltiorrhiza*) to 26,583 (*S. dabiieshanensis*). Among these codons, leucine (10.59–10.64%) was the most frequently observed amino acid, and cysteine (1.10–1.12%) was the least frequently observed (Figure 2, Table S1). Most protein-coding genes had the same standard ATG sequence as the initiation codon, but the *rps19* and *ycf15* genes started with a GTG sequence. This non-ATG initiation codon has also been found in other angiosperm chloroplasts [28,57,58]. Heatmaps showed that 30 of the total 64 types of codons used in sect. *Drymosphace* were preferred codons (RSCU > 1), and with the exception of UUG (Leu), all of the preferred codons ended with an A or a U (Figure 3, Supplementary Material Table S1). This codon usage bias shows a similar trend observed in the majority of plastomes in higher plants [52,59,60].

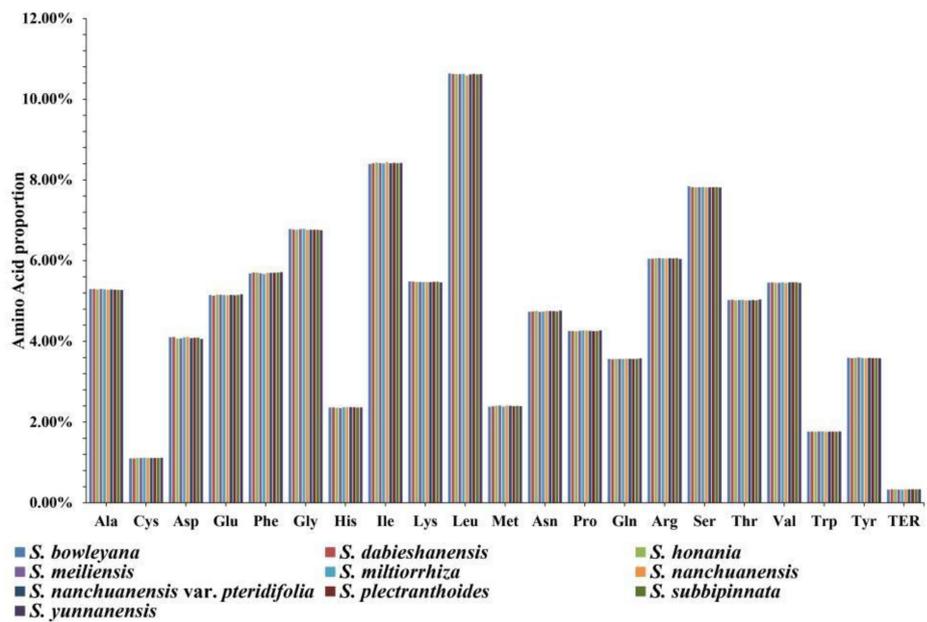


Figure 2. Amino acid frequencies in all protein-coding genes in the chloroplast genome of 10 taxa of *Salvia* sect. *Drymospace*.

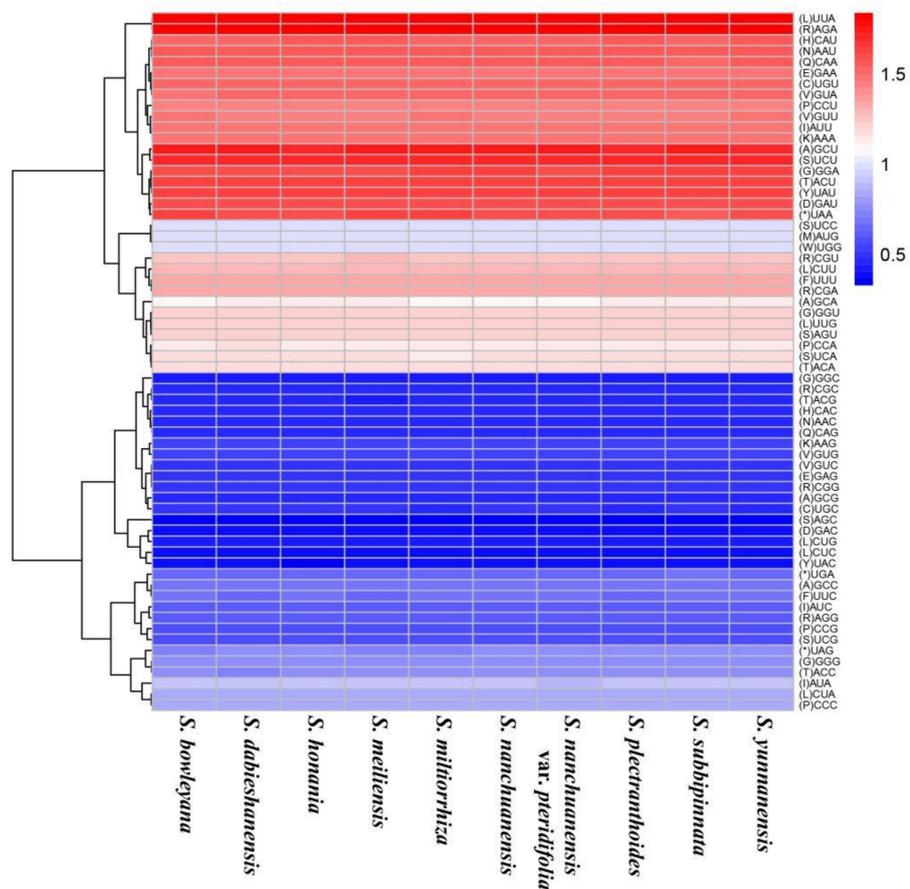


Figure 3. Heatmap of codon distribution of all protein-coding genes from 10 taxa of *Salvia* sect. *Drymospace*. Higher red values indicate higher RSCU values, and lower blue values indicate lower RSCU values.

3.3. Simple Sequence Repeats and Long Repeat Sequences

Simple sequence repeat (SSR), also called microsatellite, is a sequence with a tandemly repeated motif that ranges in size from one to six. Due to high repeatability and polymorphism, SSR is a molecular marker that is commonly used in population genetics and evolutionary analysis [52,61–64]. In this study, a total of 1685 SSRs were identified with each taxon bearing 166–173 SSRs from 10 plastomes of sect. *Drymosphace* (Figure 4, Table S2). Five types of SSR (excluding penta-nucleotide SSR) were detected in sect. *Drymosphace*, of which the mono-nucleotide SSR was the richest (1265, 75.07%), followed by the di-nucleotide (346, 20.53%), tetra-nucleotide (69, 4.09%), hexa-nucleotide (4, 0.24%), and tri-nucleotide (1, 0.06%) SSRs. In terms of repeat units, the A/T unit in the mono-nucleotide repeat was the most abundant (1242, 73.71%), followed by the di-nucleotide AT/AT (136, 8.07%) and TA/TA (100, 5.93%). The tri-nucleotide SSR (AAT/ATT) was only detected once (1, 0.06%) in *S. nanchuanensis* var. *pteridifolia*, and the hexa-nucleotide SSRs (4, 0.24%) were detected once in *S. dabieshanensis* (ATTCAT/ATGAAT), *S. meiliensis* (AATCAA/TTGATT), *S. miltiorrhiza* (ACTTAG/CTAAGT), and *S. plectranthoides* (ACTTAG/CTAAGT). In addition, six unique repeat units (AAT/ATT, AAAT/ATTT, TTAA/AATT, ATTCAT/ATGAAT, AATCAA/TTGATT, ACTTAG/CTAAGT) were detected within sect. *Drymosphace*. The majority of the SSRs resided in LSC regions (1125, 66.77%), and the others were located in other SSC regions (320, 18.99%) and in IR regions (240, 14.24%). As reported in previous plastomes of *Salvia* of Lamiaceae [28,51] as well as in other families [55,65,66], most of the SSRs in sect. *Drymosphace* are composed of A/T mono-nucleotides, which represent a percentage of 73.71%. Similar to most *Salvia* plastomes that have been reported upon before [28,51,67], we did not detect the penta-nucleotide SSR in sect. *Drymosphace*. However, a recent study reported a penta-nucleotide SSR from *S. yunnanensis* of sect. *Drymosphace* [68]. *Salvia yunnanensis* was also included in this study as well as in another independent analysis [28], in which the samples were from Guizhou and Yunnan Province, China, respectively, but neither study identified the penta-nucleotide SSR in these species. Therefore, there is a lack of the penta-nucleotide within sect. *Drymosphace*, or at the very least, it is rare. Even though the mono-nucleotide SSRs showed significant differences among species, ranging from 123 (*S. miltiorrhiza*) to 134 (*S. yunnanensis*), no significant differences were observed in the number of the other four types of SSRs.

A total of 478 long repeat sequences with lengths greater than 30 bp, including 222 forward and 255 palindromic repeats and one complementary repeat, were detected in sect. *Drymosphace* (Figure 5, Table S3). The number of repeats for each taxa varied from 42 to 49 and had lengths that ranged from 30 to 99 bp. Most of the long repeats were located in IR (389, 81.38%) regions, followed by LSC (79, 16.53%) and SSC (10, 2.09%). In particular, most of these repeats were detected in coding regions (378, 79.08%), with a few being located in intergenic regions (50, 10.46%) and in introns (50, 10.46%). Long repeat sequences may play a pivotal role in plant evolution and could promote variation and rearrangement in the plastid genome [69–71]. All of these repeats, together with the aforementioned SSRs, may have potential utility in population studies.

3.4. Comparative Genomic Analyses

The contraction and expansion of IR regions at the borders is a main reason for size variation in plastomes and plays an important role in the evolution of seed plants [72–75]. To explore the expansion and contraction of the IR regions, the boundaries between the SC and IR regions of sect. *Drymosphace* were analyzed. The *rps19* gene spanned the LSC/IRb boundary in all species and had a length of 234 bp or 236 bp in the LSC and 43 bp or 45 bp in the IRb, which generated a short *rps19* pseudogene ($\psi rps19$) that was 43 bp or 45 bp in length at the IRa/LSC border. The *ndhF* crossed the IRb/SSC border and had an equal length of 32 bp in the IRb and of 2185 bp or 2203 bp in the SSC. The SSC/IRa boundary was located within the *ycf1* gene and had lengths ranging from 4452 bp to 4470 bp in the SSC and of 1056 bp or 1167 bp in the IRa, resulting in a pseudogene ($\psi ycf1$) at the IRb/SSC border with 1056 bp or 1167 bp overlapping with the *ndhF* gene. In addition, the IRa/LSC

boundary resided between the *ψrps19* and *trnH*, and the distance from *trnH* was 8 bp or 10 bp (Figure 6). In *Salvia*, IR/LSC junctions are considered to be highly conserved, and the expansion/contraction of the IR regions is infrequent [28]. Although a few IR contractions have been reported in *Salvia* (e.g., *S. hispanica*, *S. mekongensis*, and *S. rosmarinus*) [28], no contractions were detected in this study, suggesting that IR contraction events may not occur in sect. *Drymosphace*.

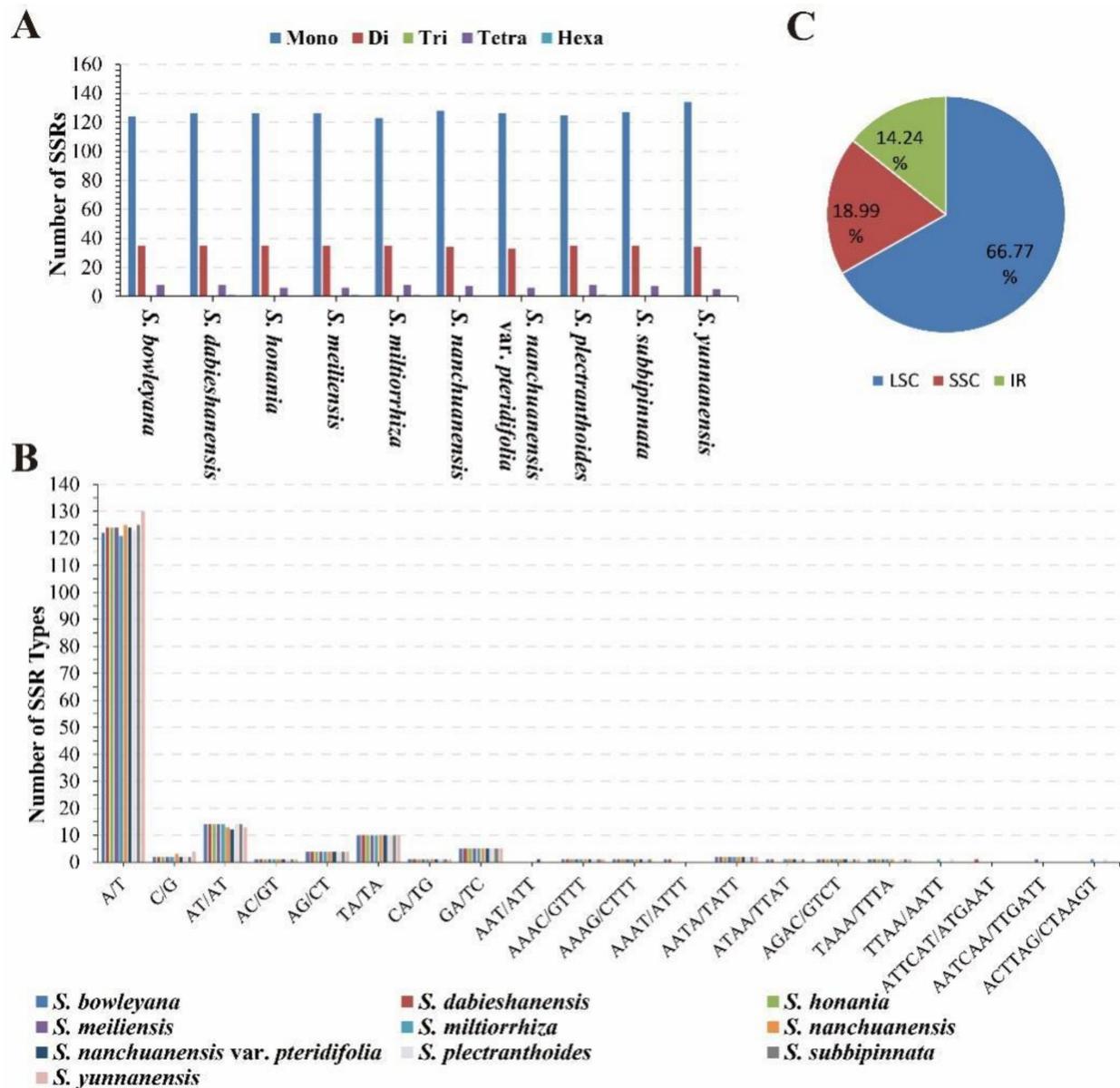


Figure 4. Comparison of simple sequence repeats (SSR) among the plastomes of 10 taxa of *Salvia* sect. *Drymosphace*. (A). Number of different types of SSRs; (B). number of different of SSR repeat units; (C) frequencies of SSRs in LSC, IR, and SSC regions.

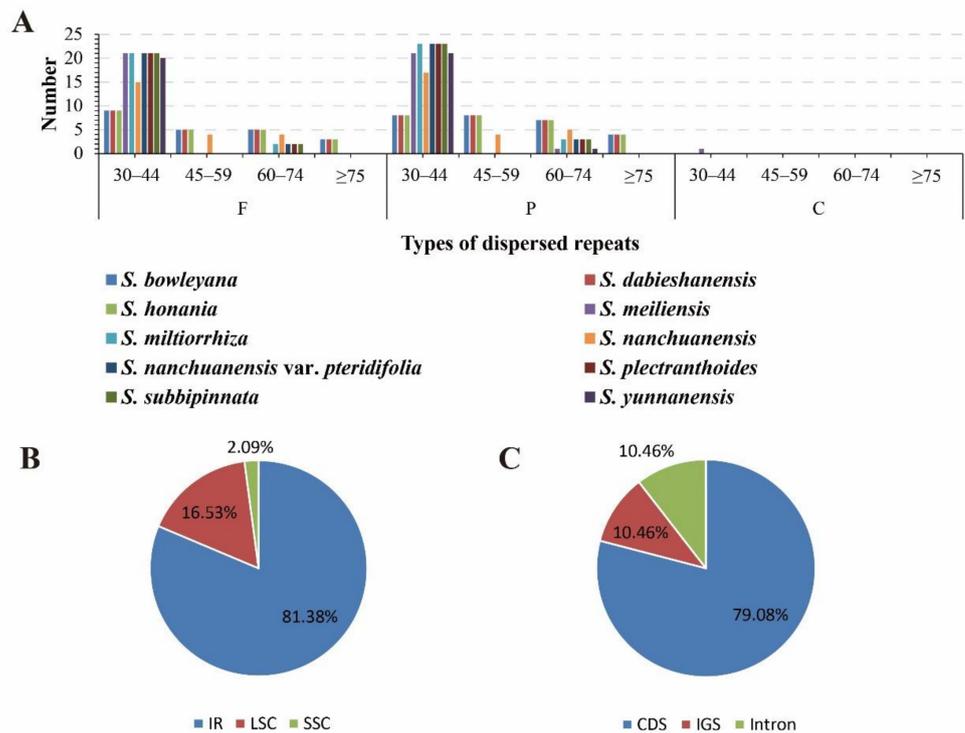


Figure 5. Analysis of repetitive sequences in the plastomes of 10 taxa of *Salvia* sect. *Drymosphace*. (A). Number of different types of longer repeats; (B). frequencies of longer repeats in the LSC, SSC, and IR regions; (C). frequencies of longer repeats in protein coding regions, intergenic spacers, and intron regions.

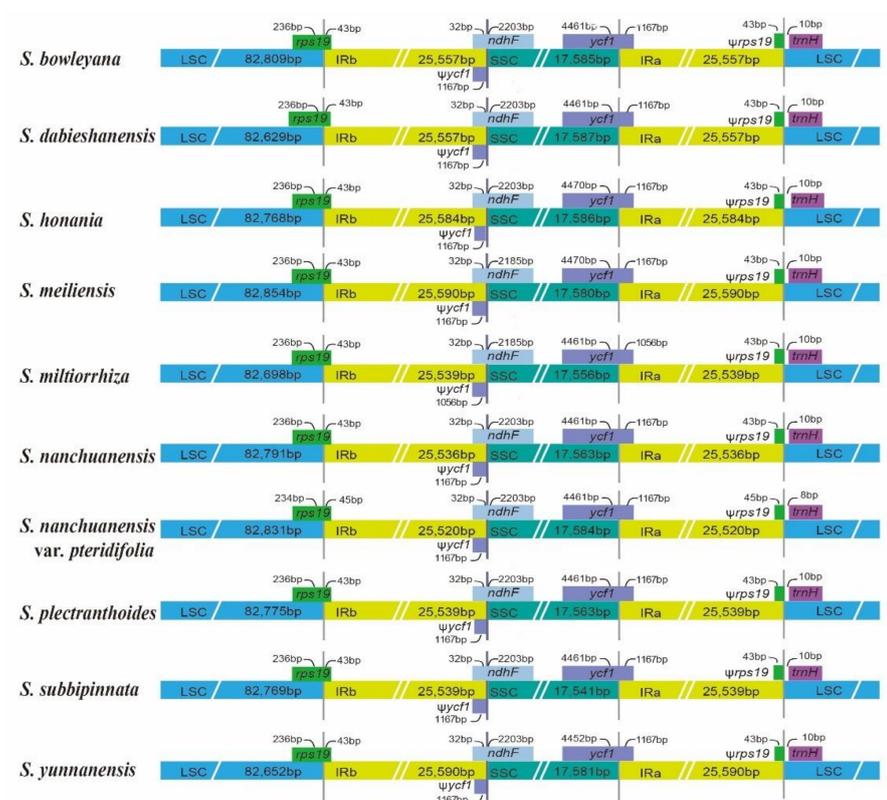


Figure 6. Comparison of junctions between the LSC, IR, and SSC regions among plastomes of 10 taxa of *Salvia* sect. *Drymosphace*. Genes are denoted by colored boxes, with ψ indicating a pseudogene.

Comparison of complete plastome sequences showed that the genes within sect. *Drymosphace* were arranged in the same order and that the IR regions were more conserved than the LSC and SSC regions (Figure 7), which may be related to the copy corrections caused by gene conversion between the two IR regions [76,77]. In addition, noncoding regions were found to be more divergent than coding regions, and the most varied regions occurred in intergenic regions, such as *accD-psaI*, *rps4-trnL*, *rps16-trnQ*, and *trnH-psbA*. Genome rearrangement analysis showed a collinear relationship, suggesting that there was no rearrangement or inversion events in the plastomes of sect. *Drymosphace* (Figure S1).

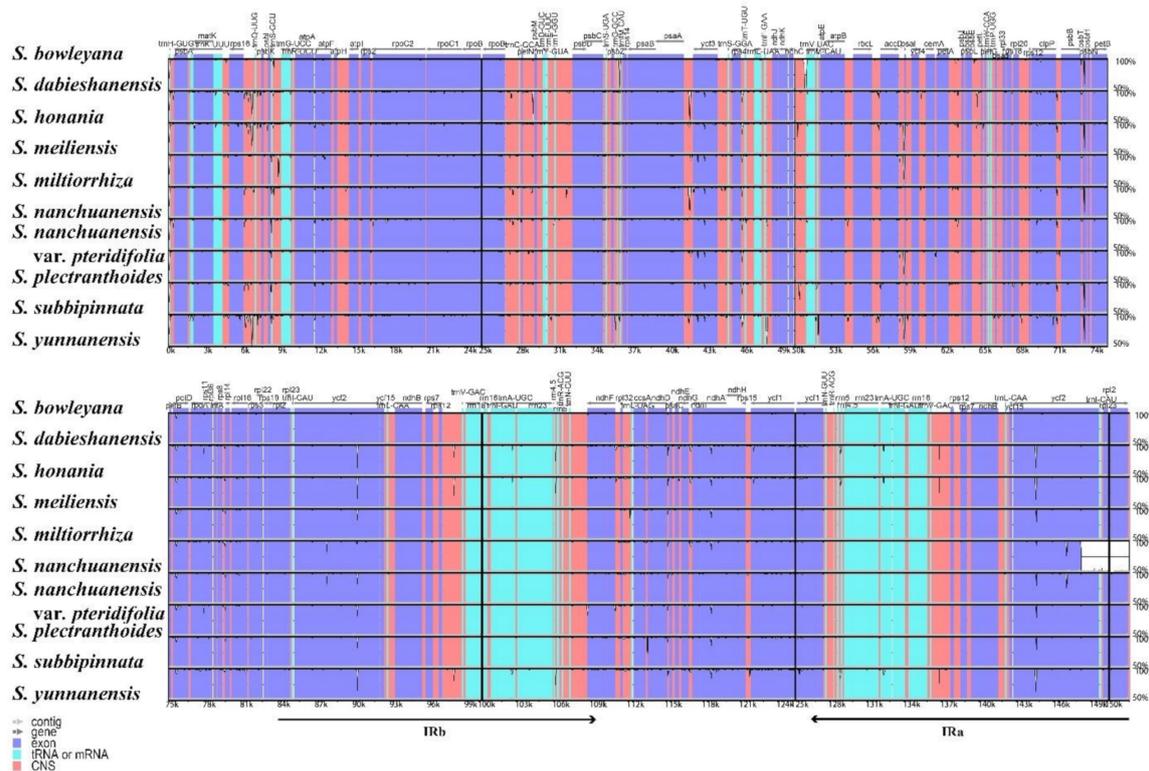


Figure 7. Sequence alignment of the whole plastomes of 10 taxa of *Salvia* sect. *Drymosphace* using mVISTA with *S. bowleyana* as a reference. The vertical scale indicates percentage of identity, ranging from 50% to 100%.

Sequence divergence analyses showed that the nucleotide variability (Pi) values of the sect. *Drymosphace* plastomes ranged from 0 to 0.01376, with an average value of 0.0014 (Figure 8). Nine highly variable regions with Pi values > 0.005 were detected, including two genes (*clpP* and *ycf1*) and seven intergenic spacers (*trnH-psbA*, *trnK-rps16*, *petN-psbM*, *rps4-trnT*, *rbcL-accD*, *rpl32-trnL-ccsA*, and *ndhD-psaC*). Among these hyper-variable regions, six loci were located in the LSC, three were located in the SSC, and none were located in the IR regions. *trnH-psbA* is the most variable region (Pi = 0.01376) among the nine hypervariable regions. Due to high rates of insertion/deletion and universal primers, *trnH-psbA* has been selected to be a plant barcode for species discrimination [78–81] and has been used for phylogenetic analyses in Lamiaceae [82–84]. However, the resolution of this DNA marker is not high enough to solve phylogenetic relationships at the section and below levels for *Salvia* [2,85]. For the two hypervariable genic regions, the *clpP* gene has not been used as a marker in phylogenetic study of *Salvia*, and *ycf1* has been demonstrated to be a promising maker for phylogenetic analyses within *Salvia* [28,86].

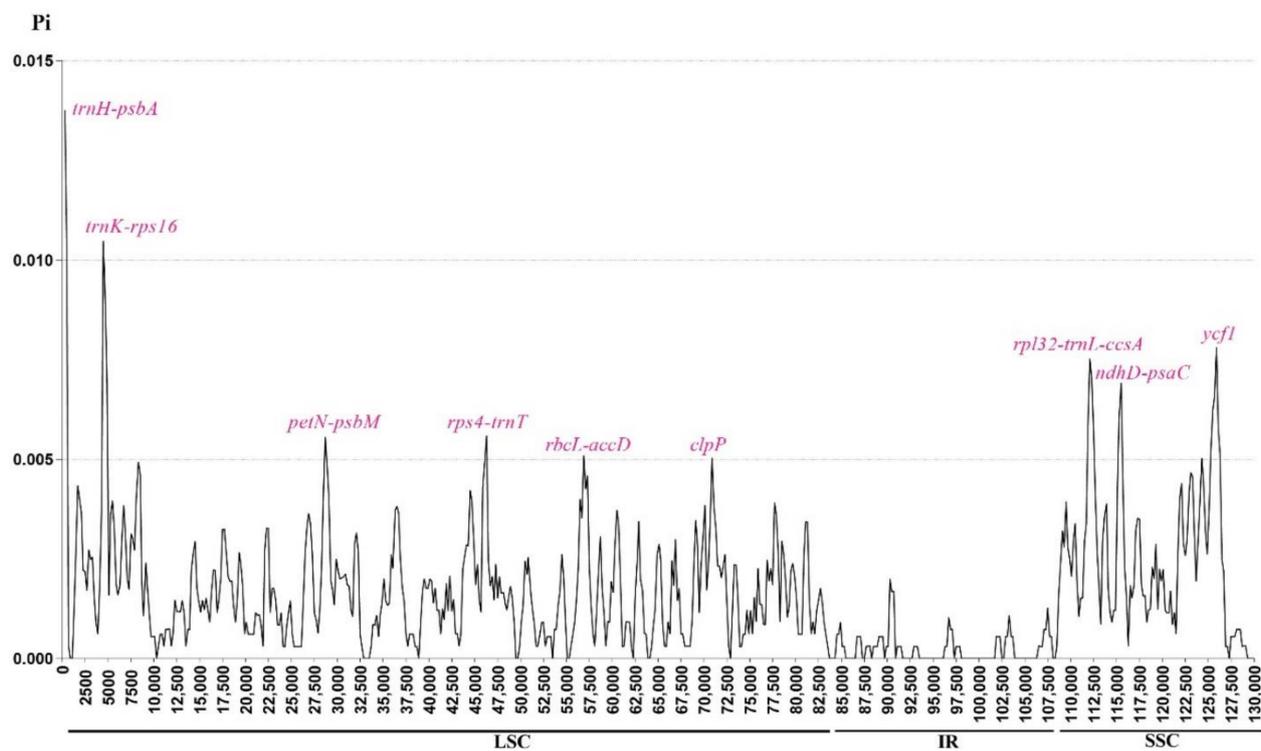


Figure 8. Sliding window analysis of plastomes of 10 taxa of *Salvia* sect. *Drymosphace*. The X-axis indicates the position of the midpoint of a window, and the Y-axis represents the nucleotide diversity of each window. The nine hyper-variable regions with $Pi > 0.005$ are marked.

3.5. Phylogenetic Analysis

Seven matrices including 17 taxa were employed to conduct phylogenetic analyses (Table 1). The aligned CP dataset had the longest length (145,835 bp), and the combined nine HVRs had the shortest length (10,298 bp). The highest percentage of constant sites were from the IR data set (99.11%), and the lowest were from the HVR data set (88.33%). As opposed to the constant sites, the percentage of the informative sites in the HVR data set was the highest (3.52%), and lowest percentage was in the IR regions (0.31%) (Table 2).

The topologies that were inferred by ML and BI analyses for each matrix were identical, with the exception of a few weakly supported nodes. Therefore, only an ML tree with posterior probabilities (PP) indicated above the branches and the ML bootstrap (BS) values provided below.

In all of the analyses, monophylies of subg. *Glutinaria* and sect. *Drymosphace* were recovered, and sister relationships between *Salvia honania* and *S. meiliensis*, *S. dabieshanensis*, and *S. bowleyana* within sect. *Drymosphace* were strongly supported (Figure 9 and Figures S2–S7). Within sect. *Drymosphace*, with the exception of the CR matrix, all of the analyses supported the finding that *S. honania* and *S. meiliensis* together to be sister to the rest of the section; with the exception of the IR regions, all of the topologies that were inferred from other data set showed that *S. dabieshanensis*, *S. bowleyana*, *S. miltiorrhiza*, and *S. plectranthoides* formed a strongly supported subclade. The phylogenetic positions of *S. nanchuanensis*, *S. yunnanensis*, and *S. subbipinnata* varied slightly in different data set.

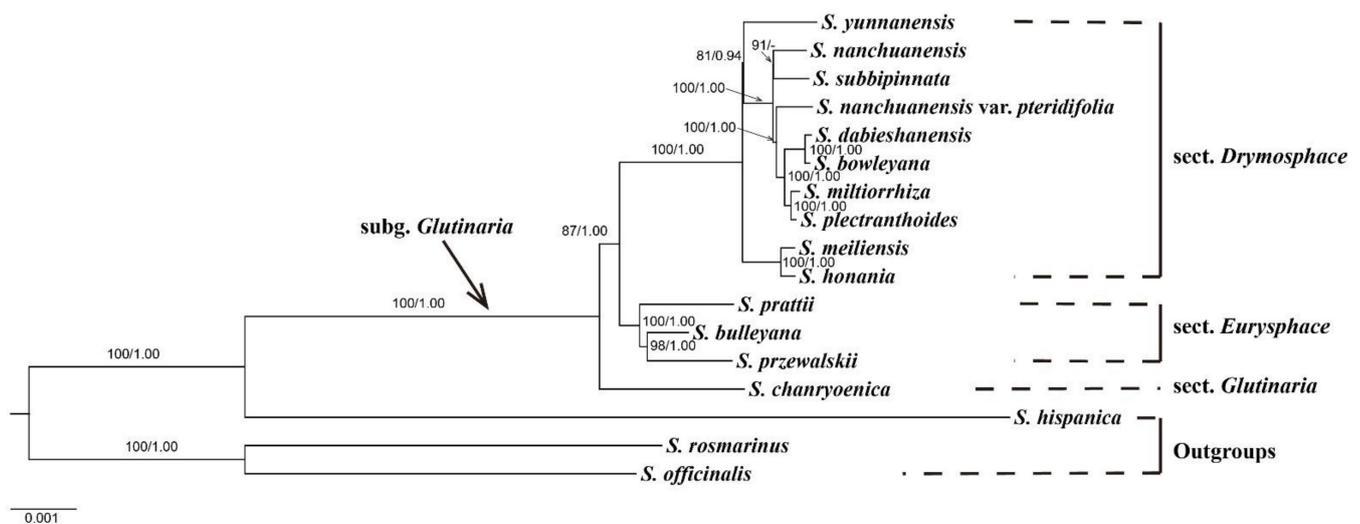


Figure 9. Phylogenetic tree inferred from maximum likelihood and Bayesian inference based on complete chloroplast genome. Posterior probabilities are provided above the branches, and the ML bootstrap values are indicated below, with the ML bootstrap values <50% and PP < 0.90 indicated by ‘-’.

Integrating molecular phylogenetic with morphological evidence, Hu et al. [2] established a new sect. *Drymosphace*, but the interspecific relationships within the section are confused, and the supported values for most subclades are weak. For example, *S. honania* and *S. meiliensis* are two morphologically similar species [87]. They are easily distinguished from the other species within sect. *Drymosphace*, as characterized by small and oblong upper corolla lips and clearly exerted stamens that adhere laterally to the corolla [2,87]. The two species should have been sister group. However, they did not gather together in previous analyses that were based on only a few DNA markers [87]. In this study, all of the analyses indicated that the two species formed a sister relationship, supporting the finding that *S. honania* is closely related to *S. meiliensis* [2,87]. Additionally, compared to previous studies including sect. *Drymosphace* based on short DNA fragments [2,12,85], the plastid phylogenomic analyses in this study based on all of the matrices greatly improved the phylogenetic resolution (Figure 9). Although monophyly of sect. *Drymosphace* and sister relationships of *S. honania* and *S. meiliensis* were confirmed in this study, the relationships between some of the taxa presented here are inconsistent with those inferred by morphological characteristics. Morphologically, *Salvia plectranthoides*, *S. nanchuanensis*, and *S. nanchuanensis* var. *pteridifolia* share long tubular corolla tubes and straight upper corolla lips and are therefore placed in ser. *Plectranthoidites* [4,16]. They should have been found to be closely related and grouped together in phylogenetic trees, especially *S. nanchuanensis* and *S. nanchuanensis* var. *pteridifolia*. However, in this study, *S. nanchuanensis* is sister to *S. subbipinnata* instead of *S. nanchuanensis* var. *pteridifolia*, and *S. plectranthoides* is sister to *S. miltiorrhiza*, indicating that *S. nanchuanensis* and ser. *Plectranthoidites* are not monophyletic. Therefore, the plastid genome is still unable to completely determine the interspecific relationships of sect. *Drymosphace*.

Of all seven data sets, the IR regions generated the poorest phylogenetic resolution, with four subclades having very low support values within sect. *Drymosphace* (PP < 0.9 and BS < 50%) (Figure S4). Wu et al. [27] also reported similar results in plastid phylogenomic results for *Salvia* [27]. The sequence length as well as informative sites may lead to this result. Previous studies indicate that more and longer DNA sequences may greatly improve the phylogenetic resolution and the support values of the branches [28,88,89]. In this study, the IR regions (25,563 bp) and HVR (10,298) were the two shortest data sets, so the phylogenetic resolutions that were inferred from them are lower than those of the other five matrices (Figures S4 and S7). On the other hand, the IR region has the lowest percentage of

informative sites, which may be the result of the low resolution. Therefore, the use of IR regions separately is not suggested for phylogenetic analysis in *Salvia*.

4. Conclusions

In this study, we sequenced and assembled eight *Salvia* sect. *Drymosphace* plastomes. The reported plastid genomes are conservative, showing high levels of consistency and similarity to the chloroplast genomes of other species of Lamiaceae in terms of gene content, order, and structure. Comparative analyses revealed that there is no rearrangement or inversion events in the plastomes of sect. *Drymosphace* and that single-copy regions and intergenic regions are more variable than IR regions and coding regions. Phylogenomic analyses can recover the monophyly of the newly established sect. *Drymosphace*, considerably improve phylogenetic resolution, and determine interspecific relationships for some of the species within this section.

Supplementary Materials: The following are available at <https://www.mdpi.com/article/10.3390/d14050324/s1>. Table S1: Amino acid frequencies and codon distribution of 10 taxa of *Salvia* sect. *Drymosphace*, Table S2: SSRs of 10 taxa of *Salvia* sect. *Drymosphace*. Table S3: Long repeats of 10 taxa of *Salvia* sect. *Drymosphace*, Figure S1: Mauve alignment of plastomes of 10 taxa of *Salvia* sect. *Drymosphace*, Figure S2: Phylogenetic tree inferred from maximum likelihood and Bayesian inference based on LSC, Figure S3: Phylogenetic tree inferred from maximum likelihood and Bayesian inference based on SSC, Figure S4: Phylogenetic tree inferred from maximum likelihood and Bayesian inference based on IR, Figure S5: Phylogenetic tree inferred from maximum likelihood and Bayesian inference based on CR, Figure S6: Phylogenetic tree inferred from maximum likelihood and Bayesian inference based on NCR, Figure S7: Phylogenetic tree inferred from maximum likelihood and Bayesian inference based on HVR.

Author Contributions: Conceptualization, G.-X.H.; software, T.S., L.G. and F.Z.; formal analysis, T.S., M.W. and L.G.; writing—original draft preparation, T.S. and G.-X.H.; writing—review and editing, G.-X.H., C.-L.X., T.S., Y.-F.G. and F.Z.; funding acquisition, G.-X.H. and Y.-F.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (32060048 and 31900275) and the Construction Program of Biology First-Class Discipline in Guizhou (GNYL [2017] 009).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The plastome sequence data supporting this study are openly available in the GenBank nucleotide database and the Sequence Read Archive. Other data are contained within the article or within the Supplementary Materials.

Acknowledgments: We thank Jia-Xin Yang, Xuanze He, and Bo Pan for their help with the field work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Drew, B.; González-Gallegos, J.G.; Xiang, C.L.; Kriebel, R.; Drummond, C.; Walker, J.; Sytsma, K. *Salvia* united: The greatest good for the greatest number. *Taxon* **2017**, *66*, 133–145. [[CrossRef](#)]
2. Hu, G.X.; Takano, A.; Drew, B.T.; Liu, E.D.; Soltis, D.E.; Soltis, P.S.; Peng, H.; Xiang, C.L. Phylogeny and staminal evolution of *Salvia* (Lamiaceae, Nepetoideae) in East Asia. *Ann. Bot.* **2018**, *122*, 649–668. [[CrossRef](#)] [[PubMed](#)]
3. Kriebel, R.; Drew, B.T.; Drummond, C.P.; Gonzalez-Gallegos, J.G.; Celep, F.; Mahdjoub, M.M.; Rose, J.P.; Xiang, C.L.; Hu, G.X.; Walker, J.B.; et al. Tracking temporal shifts in area, biomes, and pollinators in the radiation of *Salvia* (sages) across continents: Leveraging anchored hybrid enrichment and targeted sequence data. *Am. J. Bot.* **2019**, *106*, 573–597. [[CrossRef](#)]
4. Li, H.W.; Hedge, I.C. *Salvia*. In *Flora of China*; Wu, C.Y., Raven, P.H., Hong, D.Y., Eds.; Science Press: Beijing, China; Missouri Botanical Garden Press: St. Louis, MO, USA, 1994; Volume 17, pp. 196–224.
5. Hu, G.X.; Liu, Y.; Xu, W.B.; Liu, E.D. *Salvia petrophila* sp. nov. (Lamiaceae) from north Guangxi and south Guizhou, China. *Nord. J. Bot.* **2014**, *32*, 190–195. [[CrossRef](#)]

6. Hu, G.X.; Liu, E.D.; Zhang, T.; Cai, J.; Xiang, C.L. *Salvia luteistriata* (Lamiaceae), a new species from northeastern Sichuan, China. *Phytotaxa* **2017**, *314*, 123–128. [[CrossRef](#)]
7. Hu, G.X.; Peng, H. Identity of *Salvia weihaiensis* (Lamiaceae) from China. *Phytotaxa* **2015**, *202*, 298–300. [[CrossRef](#)]
8. Xiang, C.L.; Hu, G.X.; Peng, H. New combinations and new synonyms in Lamiaceae from China. *Biodivers. Sci.* **2016**, *24*, 719–722. [[CrossRef](#)]
9. Ding, B.Y.; Chen, Z.H.; Xu, Y.L.; Jin, X.F.; Wu, D.F.; Chen, J.B.; Wu, W.J. New species and combination of Lamiaceae from Zhejiang China. *Guihaia* **2019**, *39*, 10–15. [[CrossRef](#)]
10. Wei, Y.K.; Pendry, C.A.; Zhang, D.G.; Huang, Y.B. *Salvia daiguii* (Lamiaceae): A new species from west Hunan, China. *Edinb. J. Bot.* **2019**, *76*, 359–368. [[CrossRef](#)]
11. Wei, Y.K.; Pendry, C.A.; Huang, Y.B.; Ge, B.J.; Xiao, H.W. *Salvia subviolacea*, A new species from the Himalayas-Hengduan Mountains, China. *Edinb. J. Bot.* **2021**, *78*, 1–9. [[CrossRef](#)]
12. Hu, G.X.; Liu, E.D.; Wu, Z.K.; Sytsma, K.J.; Drew, B.T.; Xiang, C.L. Integrating DNA sequences with morphological analysis clarifies phylogenetic position of *Salvia grandifolia* (Lamiaceae): An enigmatic species endemic to Southwestern China. *Int. J. Plant Sci.* **2020**, *181*, 787–799. [[CrossRef](#)]
13. Bentham, G. *Labiatarum Genera Et Species*; Ridgeway: London, UK, 1932–1936.
14. Bentham, G. Labiatae. In *Genera Plantarum*; Bentham, G., Hooker, J.D., Eds.; Reeve and Co.: London, UK, 1876; Volume 2, pp. 1160–1196.
15. Stibal, E. A revision of Chinese and East Burmese species of *Salvia*. *Acta Horti Gotobg.* **1934**, *9*, 101–165.
16. Wu, C.Y. *Salvia*. In *Flora Reipublicae Popularis Sinicae*; Wu, C.Y., Ed.; Science Press: Beijing, China, 1977; Volume 66, pp. 70–196.
17. Wang, B.Q. *Salvia miltiorrhiza*: Chemical and pharmacological review of a medicinal plant. *J. Med. Plant. Res.* **2010**, *4*, 2813–2820. [[CrossRef](#)]
18. Walker, J.B.; Sytsma, K.J. Staminal evolution in the genus *Salvia* (Lamiaceae): Molecular phylogenetic evidence for multiple origins of the staminal lever. *Ann. Bot.* **2007**, *100*, 375–391. [[CrossRef](#)]
19. Jenks, A.A.; Walker, J.B.; Kim, S.C. Phylogeny of New World *Salvia* subgenus *Calosphace* (Lamiaceae) based on cpDNA (psbA-trnH) and nrDNA (ITS) sequence data. *J. Plant Res.* **2013**, *126*, 483–496. [[CrossRef](#)]
20. Will, M.; Classen-Bockhoff, R. Why Africa matters: Evolution of Old World *Salvia* (Lamiaceae) in Africa. *Ann. Bot.* **2014**, *114*, 61–83. [[CrossRef](#)]
21. Will, M.; Classen-Bockhoff, R. Time to split *Salvia* s.l. (Lamiaceae)-New insights from Old World *Salvia* phylogeny. *Mol. Phylogenet. Evol.* **2017**, *109*, 33–58. [[CrossRef](#)]
22. Fragoso-Martínez, I.; Martínez-Gordillo, M.; Salazar, G.A.; Sazatornil, F.; Jenks, A.A.; Peña, M.R.G.; Barrera-Aveleida, G.; Benitez-Vieyra, S.; Magallón, S.; Cornejo-Tenorio, G. Phylogeny of the Neotropical sages (*Salvia* subg. *Calosphace*; Lamiaceae) and insights into pollinator and area shifts. *Plant Syst. Evol.* **2018**, *304*, 43–55. [[CrossRef](#)]
23. Deng, T.; Nie, Z.L.; Drew, B.T.; Volis, S.; Kim, C.; Xiang, C.L.; Zhang, J.W.; Wang, Y.H.; Sun, H. Does the Arcto-Tertiary biogeographic hypothesis explain the disjunct distribution of Northern Hemisphere herbaceous plants? The case of *Meehania* (Lamiaceae). *PLoS ONE* **2015**, *10*, e0117171. [[CrossRef](#)]
24. Moore, M.J.; Bell, C.D.; Soltis, P.S.; Soltis, D.E. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19363–19368. [[CrossRef](#)]
25. Parks, M.; Cronn, R.; Liston, A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.* **2009**, *7*, 84. [[CrossRef](#)] [[PubMed](#)]
26. Zhao, F.; Li, B.; Drew, B.T.; Chen, Y.P.; Wang, Q.; Yu, W.B.; Liu, E.D.; Salmaki, Y.; Peng, H.; Xiang, C.L. Leveraging plastomes for comparative analysis and phylogenomic inference within *Scutellarioideae* (Lamiaceae). *PLoS ONE* **2020**, *15*, e0232602. [[CrossRef](#)] [[PubMed](#)]
27. Wu, H.; Ma, P.F.; Li, H.T.; Hu, G.X.; Li, D.Z. Comparative plastomic analysis and insights into the phylogeny of *Salvia* (Lamiaceae). *Plant Divers.* **2021**, *43*, 15–26. [[CrossRef](#)] [[PubMed](#)]
28. Zhao, F.; Drew, B.T.; Chen, Y.P.; Hu, G.X.; Li, B.; Xiang, C.L. The Chloroplast genome of *Salvia*: Genomic characterization and phylogenetic analysis. *Int. J. Plant Sci.* **2020**, *181*, 812–830. [[CrossRef](#)]
29. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
30. Jin, J.J.; Yu, W.B.; Yang, J.B.; Song, Y.; De Pamphilis, C.W.; Yi, T.S.; Li, D.Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **2020**, *21*, 241. [[CrossRef](#)]
31. Wick, R.R.; Schultz, M.B.; Zobel, J.; Holt, K.E. Bandage: Interactive visualization of de novo genome assemblies. *Bioinformatics* **2015**, *31*, 3350–3352. [[CrossRef](#)]
32. Qu, X.J.; Moore, M.J.; Li, D.Z.; Yi, T.S. PGA: A software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* **2019**, *15*, 50. [[CrossRef](#)]
33. Chen, H.; Zhang, J.; Yuan, G.; Liu, C. Complex interplay among DNA modification, noncoding RNA expression and protein-coding RNA expression in *Salvia miltiorrhiza* chloroplast genome. *PLoS ONE* **2014**, *9*, e99314. [[CrossRef](#)]
34. Lowe, T.M.; Chan, P.P. tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* **2016**, *44*, W54–W57. [[CrossRef](#)]

35. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)] [[PubMed](#)]
36. Lohse, M.; Drechsel, O.; Kahlau, S.; Bock, R. OrganellarGenomeDRAW—A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **2013**, *41*, W575–W581. [[CrossRef](#)] [[PubMed](#)]
37. Peden, J.F. Analysis of Codon Usage. Ph.D. Thesis, University of Nottingham, Nottingham, UK, 1999.
38. Randall, T.A.; Dwyer, R.A.; Huitema, E.; Beyer, K.; Cvitanich, C.; Kelkar, H.; Fong, A.M.; Gates, K.; Roberts, S.; Yatzkan, E.; et al. Large-scale gene discovery in the Oomycete *Phytophthora infestans* reveals likely components of phytopathogenicity shared with true fungi. *Mol. Plant Microbe Interact.* **2005**, *18*, 229–243. [[CrossRef](#)]
39. Thiel, T.; Michalek, W.; Varshney, R.K.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [[CrossRef](#)] [[PubMed](#)]
40. 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](#)]
41. Mayor, C.; Brudno, M.; Schwartz, J.R.; Poliakov, A.; Rubin, E.M.; Frazer, K.A.; Pachter, L.S.; Dubchak, I. VISTA: Visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **2000**, *16*, 1046–1047. [[CrossRef](#)] [[PubMed](#)]
42. Kurtz, S.; Phillippy, A.; Delcher, A.L.; Smoot, M.; Shumway, M.; Antonescu, C.; Salzberg, S.L. Versatile and open software for comparing large genomes. *Genome Biol.* **2004**, *5*, R12. [[CrossRef](#)]
43. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [[CrossRef](#)]
44. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
45. Talavera, G.; Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **2007**, *56*, 564–577. [[CrossRef](#)]
46. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]
47. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 14 November 2010; IEEE: Manhattan, NY, USA, 2010. [[CrossRef](#)]
48. 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](#)] [[PubMed](#)]
49. Zhang, D.; Gao, F.L.; Jakovlic, 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](#)]
50. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermini, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [[CrossRef](#)]
51. Qian, J.; Song, J.; Gao, H.; Zhu, Y.; Xu, J.; Pang, X.; Yao, H.; Sun, C.; Li, X.; Li, C.; et al. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. *PLoS ONE* **2013**, *8*, e57607. [[CrossRef](#)]
52. Gao, B.; Yuan, L.; Tang, T.; Hou, J.; Pan, K.; Wei, N. The complete chloroplast genome sequence of *Alpinia oxyphylla* Miq. and comparison analysis within the Zingiberaceae family. *PLoS ONE* **2019**, *14*, e0218817. [[CrossRef](#)] [[PubMed](#)]
53. Zong, D.; Zhou, A.; Zhang, Y.; Zou, X.; Li, D.; Duan, A.; He, C. Characterization of the complete chloroplast genomes of five *Populus* species from the western Sichuan plateau, southwest China: Comparative and phylogenetic analyses. *PeerJ* **2019**, *7*, e6386. [[CrossRef](#)] [[PubMed](#)]
54. Yang, J.X.; Hu, G.X.; Hu, G.W. Comparative genomics and phylogenetic relationships of two endemic and endangered species (*Handeliidendron bodinieri* and *Eurycorymbus cavaleriei*) of two monotypic genera within Sapindales. *BMC Genom.* **2022**, *23*, 27. [[CrossRef](#)]
55. Gu, L.; Su, T.; An, M.T.; Hu, G.X. The Complete chloroplast genome of the vulnerable *Oreocharis esquirolii* (Gesneriaceae): Structural features, comparative and phylogenetic Analysis. *Plants* **2020**, *9*, 1692. [[CrossRef](#)]
56. 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](#)]
57. Millen, R.S.; Olmstead, R.G.; Adams, K.L.; Palmer, J.D.; Lao, N.T.; Heggie, L.; Kavanagh, T.A.; Hibberd, J.M.; Gray, J.C.; Morden, C.W.; et al. Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell* **2001**, *13*, 645–658. [[CrossRef](#)] [[PubMed](#)]
58. Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.; Matsubayashi, T.; Zaita, N.; Chunwongse, J.; Obokata, J.; Yamaguchi-Shinohara, K.; et al. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **1986**, *5*, 2043–2049. [[CrossRef](#)]
59. Shimada, H.; Sugiura, M. Fine structural features of the chloroplast genome: Comparison of the sequenced chloroplast genomes. *Nucleic Acids Res.* **1991**, *19*, 983–995. [[CrossRef](#)] [[PubMed](#)]

60. Park, I.; Kim, W.J.; Yeo, S.M.; Choi, G.; Kang, Y.M.; Piao, R.; Moon, B.C. The complete chloroplast genome sequences of *Fritillaria ussuriensis* Maxim. and *Fritillaria cirrhosa* D. Don, and comparative analysis with Other *Fritillaria* species. *Molecules* **2017**, *22*, 982. [[CrossRef](#)] [[PubMed](#)]
61. Powell, W.; Morgante, M.; McDevitt, R.; Vendramin, G.G.; Rafalski, J.A. Polymorphic simple sequence repeat regions in chloroplast genomes: Applications to the population genetics of pines. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7759–7763. [[CrossRef](#)] [[PubMed](#)]
62. Chen, C.; Zhou, P.; Choi, Y.A.; Huang, S.; Gmitter, F.G., Jr. Mining and characterizing microsatellites from *Citrus* ESTs. *Theor. Appl. Genet.* **2006**, *112*, 1248–1257. [[CrossRef](#)] [[PubMed](#)]
63. Feng, S.; He, R.; Lu, J.; Jiang, M.; Shen, X.; Jiang, Y.; Wang, Z.; Wang, H. Development of SSR markers and assessment of genetic diversity in medicinal *Chrysanthemum morifolium* cultivars. *Front. Genet.* **2016**, *7*, 113. [[CrossRef](#)] [[PubMed](#)]
64. He, Y.; Xiao, H.; Deng, C.; Xiong, L.; Yang, J.; Peng, C. The complete chloroplast genome sequences of the medicinal plant *Pogostemon cablin*. *Int. J. Mol. Sci.* **2016**, *17*, 820. [[CrossRef](#)]
65. Kuang, D.Y.; Wu, H.; Wang, Y.L.; Gao, L.M.; Zhang, S.Z.; Lu, L. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): Implication for DNA barcoding and population genetics. *Genome* **2011**, *54*, 663–673. [[CrossRef](#)]
66. Du, Y.P.; Bi, Y.; Yang, F.P.; Zhang, M.F.; Chen, X.Q.; Xue, J.; Zhang, X.H. Complete chloroplast genome sequences of *Lilium*: Insights into evolutionary dynamics and phylogenetic analyses. *Sci. Rep.* **2017**, *7*, 5751. [[CrossRef](#)]
67. Liang, C.L.; Wang, L.; Lei, J.; Duan, B.Z.; Ma, W.S.; Xiao, S.M.; Qi, H.J.; Wang, Z.; Liu, Y.Q.; Shen, X.F.; et al. A comparative analysis of the chloroplast genomes of four *Salvia* medicinal plants. *Engineering* **2019**, *5*, 907–915. [[CrossRef](#)]
68. Liang, C.L.; Wang, L.; Ma, W.S.; Xu, J. A comparative study of complete chloroplast genome for the genus *Salvia*. *J. Plant Biochem. Biot.* **2021**, *30*, 117–125. [[CrossRef](#)]
69. Cavalier-Smith, T. Chloroplast evolution: Secondary symbiogenesis and multiple losses. *Curr. Biol.* **2002**, *12*, R62–R64. [[CrossRef](#)]
70. Wang, R.J.; Cheng, C.L.; Chang, C.C.; Wu, C.L.; Su, T.M.; Chaw, S.M. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol. Biol.* **2008**, *8*, 36. [[CrossRef](#)] [[PubMed](#)]
71. Nie, X.; Lv, S.; Zhang, Y.; Du, X.; Wang, L.; Biradar, S.S.; Tan, X.; Wan, F.; Weining, S. Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*). *PLoS ONE* **2012**, *7*, e36869. [[CrossRef](#)] [[PubMed](#)]
72. Davis, J.I.; Soreng, R.J. Migration of endpoints of two genes relative to boundaries between regions of the plastid genome in the grass family (Poaceae). *Am. J. Bot.* **2010**, *97*, 874–892. [[CrossRef](#)]
73. Li, R.; Ma, P.F.; Wen, J.; Yi, T.S. Complete sequencing of five Araliaceae chloroplast genomes and the phylogenetic implications. *PLoS ONE* **2013**, *8*, e78568. [[CrossRef](#)]
74. Menezes, A.P.A.; Resende-Moreira, L.C.; Buzatti, R.S.O.; Nazareno, A.G.; Carlsen, M.; Lobo, F.P.; Kalapothakis, E.; Lovato, M.B. Chloroplast genomes of *Byrsonima* species (Malpighiaceae): Comparative analysis and screening of high divergence sequences. *Sci. Rep.* **2018**, *8*, 2210. [[CrossRef](#)] [[PubMed](#)]
75. Shi, H.; Yang, M.; Mo, C.; Xie, W.; Liu, C.; Wu, B.; Ma, X. Complete chloroplast genomes of two *Siraitia Merrill* species: Comparative analysis, positive selection and novel molecular marker development. *PLoS ONE* **2019**, *14*, e0226865. [[CrossRef](#)]
76. Khakhlova, O.; Bock, R. Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J.* **2006**, *46*, 85–94. [[CrossRef](#)]
77. Meng, X.X.; Xian, Y.F.; Xiang, L.; Zhang, D.; Shi, Y.H.; Wu, M.L.; Dong, G.Q.; Ip, S.P.; Lin, Z.X.; Wu, L.; et al. Complete chloroplast genomes from *Sanguisorba*: Identity and variation among four species. *Molecules* **2018**, *23*, 2137. [[CrossRef](#)] [[PubMed](#)]
78. Kress, W.; Wurdack, K.J.; Zimmer, E.A.; Weigt, L.A.; Janzen, D.H. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8369–8374. [[CrossRef](#)] [[PubMed](#)]
79. Kress, W.J.; Erickson, D.L. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE* **2007**, *2*, e508. [[CrossRef](#)] [[PubMed](#)]
80. Chase, M.W.; Cowan, R.S.; Hollingsworth, P.M.; Berg, C.V.D.; Wilkinson, M.J. A Proposal for a Standardised Protocol to Barcode All Land Plants. *Taxon* **2007**, *56*, 295–299. [[CrossRef](#)]
81. Group, C.P.W. A DNA barcode for land plants. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12794–12797.
82. Meng, H.H.; Zhang, M.L. Phylogeography of *Lagochilus ilicifolius* (Lamiaceae) in relation to Quaternary climatic oscillation and aridification in northern China. *Biochem. Syst. Ecol.* **2021**, *39*, 787–796. [[CrossRef](#)]
83. Hu, G.X.; Su, T.; An, M.T.; Wang, X.Y. Rediscovery of *Pogostemon dielsianus* (Lamiaceae, Lamioideae), a rare endemic species from southwestern China, after one century. *PhytoKeys* **2021**, *171*, 61–73. [[CrossRef](#)]
84. Xiang, C.L.; Zhao, F.; Cantino, P.D.; Drew, B.T.; Li, B.; Liu, E.D.; Soltis, D.E.; Soltis, P.S.; Peng, H. Molecular systematics of *Caryopteris* (Lamiaceae) and its allies with reference to the molecular phylogeny of subfamily Ajugoideae. *Taxon* **2018**, *67*, 376–394. [[CrossRef](#)]
85. Li, Q.Q.; Li, M.H.; Yuan, Q.J.; Cui, Z.H.; Huang, L.Q.; Xiao, P.G. Phylogenetic relationships of *Salvia* (Lamiaceae) in China: Evidence from DNA sequence datasets. *J. Syst. Evol.* **2013**, *51*, 184–195. [[CrossRef](#)]
86. Drew, B.T.; Sytsma, K.J. Testing the monophyly and placement of *Lepechinia* in the tribe Mentheae (Lamiaceae). *Syst. Bot.* **2011**, *36*, 1038–1049. [[CrossRef](#)]
87. Su, S.W.; Shen, Z.A.; He, J.Q. New species of the genus *Salvia* from Anhui. *Acta Bot. Yunnanica* **1984**, *6*, 55–62.

-
88. Rokas, A.; Carroll, S.B. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Mol. Biol. Evol.* **2005**, *22*, 1337–1344. [[CrossRef](#)] [[PubMed](#)]
 89. Philippe, H.; Brinkmann, H.; Lavrov, D.V.; Littlewood, D.T.J.; Manuel, M.; Wörheide, G.; Baurain, D. Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biol.* **2011**, *9*, e1000602. [[CrossRef](#)] [[PubMed](#)]