



## Article

# Comparative Plastome Analyses of *Ephedra przewalskii* and *E. monosperma* (Ephedraceae)

Shuang Han <sup>1,2</sup> , Mingze Xia <sup>3</sup>, Jingya Yu <sup>1,2</sup>, Hao Xu <sup>1,2</sup>, Yun Han <sup>1,2</sup> and Faqi Zhang <sup>1,4,\*</sup> 

<sup>1</sup> Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology & Institute of Sanjiangyuan National Park, Chinese Academy of Sciences, Xining 810008, China

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100039, China

<sup>3</sup> School of Pharmacy, Weifang Medical University, Weifang 261053, China

<sup>4</sup> Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Xining 810008, China

\* Correspondence: fqzhang@nwipb.cas.cn; Tel.: +86-0971-6105845

**Abstract:** *Ephedra* species were erect, branching shrubs found in desert or arid regions worldwide as the source of ephedrine alkaloids. In this study, the complete chloroplast genome of *Ephedra przewalskii* and *E. monosperma* on the Qinghai-Tibet Plateau were sequenced, assembled, and annotated. Compared with the other four published *Ephedra* species, the chloroplast genomes of *Ephedra* species were highly conservative, with a quadripartite structure. The length of the chloroplast genome was 109,569 bp in *E. przewalskii* with 36.6% GC and 109,604 bp in *E. monosperma* with 36.6% GC. We detected 118 genes in both *Ephedra* species, including 73 PCGs, 37 tRNA genes, and eight rRNA genes. Among them, the *ndh* family genes were lost, which could be used to study the phylogeny and genetic diversity of the genus *Ephedra*, combined with multiple highly variable intergenic spacer (IGS) regions. Codon usage preference of *Ephedra* species was weak. The ratio of non-synonymous substitutions and synonymous substitutions was low, showing that the PCGs of *Ephedra* may be under the pressure of purifying selection. ML and BI analysis showed similar phylogenetic topologies. *Ephedra* species clustered together in a well-supported monophyletic clade. *E. przewalskii* and *E. monosperma* were not gathered in one clade, consistent with the classification system by Flora of China. This study reveals differences in the chloroplast genomes of *Ephedra*, providing valuable and abundant data for the phylogenetic analysis and species identification of *Ephedra*.

**Keywords:** *Ephedra*; chloroplast genome; phylogeny; synonymous substitutions; Qinghai-Tibet Plateau



**Citation:** Han, S.; Xia, M.; Yu, J.; Xu, H.; Han, Y.; Zhang, F. Comparative Plastome Analyses of *Ephedra przewalskii* and *E. monosperma* (Ephedraceae). *Diversity* **2022**, *14*, 792. <https://doi.org/10.3390/d14100792>

Academic Editors: Hong-Hu Meng and Jorge Capelo

Received: 25 July 2022

Accepted: 20 September 2022

Published: 24 September 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 genus *Ephedra* (Ephedraceae) is mainly distributed in desert and arid regions, with approximately 40 species worldwide. 14 species of *Ephedra* are distributed in China [1]. Potential divergence factors of *Ephedra* include the uplift of the Qinghai-Tibetan Plateau (QTP) and the Asian aridification [2]. Since Linnaeus established *Ephedra* in 1753, several scientists have held diverse views on classification system revisions within this genus [3–7].

The species of *Ephedra* are known for their ecological and medicinal values. Due to a well-developed root system with drought- and cold-resistant characteristics, it can be used in sand fixation and soil conservation programs. It has also long been an important medicinal plant in China. Containing a plethora of chemical components, it can be used to treat a variety of diseases including cold, asthma, hay fever, and urticaria [1,8]. In addition, *Ephedra* is a good source of ephedrine alkaloids that can be used to make weight-loss medicine and illicit drugs such as methamphetamine in Western countries [9]. Besides, extracts of *E. sinica* may be useful in the treatment of COVID-19. [10].

*E. przewalskii* has a much lower ephedrine content than other *Ephedra* species. However, it contains synthesis pathways for stilbene, diarylheptanoid, and other medicinal components. Stilbene has a variety of biological activities, including disease-resistance, anti-oxidation, anti-tumor, and anti-inflammatory activity. Also, the diarylheptanoid has

anti-tumor activity [11,12]. Direct ionization mass spectrometry or ITS2 barcodes are commonly used to identify *Ephedra* species [13,14]. Studies of *Ephedra* have concentrated on transcriptome data mining, the medicinal value of its chemical compounds, and the classification of morphological characteristics [11–14]. Despite numerous previous studies, it is difficult to distinguish *Ephedra* species based on morphological features [15]. However, several studies, most of which did reveal the phylogenetic relationship between the *Ephedra* species used chloroplast DNA fragments or nucleus DNA (ITS sequences) [16,17]. Also, some other studies employed complete plastid genome sequences. These studies provided new insights and ideas for dealing with the phylogenetic issues of *Ephedra* [18,19].

The chloroplast, having a small genome and being inherited uniparentally, is an essential organelle for photosynthesis [20]. Fairly conservative in their structure and sequence plastomes have evolved into an effective tool in plant evolutionary and systematic studies [21,22]. Most land plants have a chloroplast genome that is 120 to 160 kb in size. It has a quadripartite structure, consisting of two single-copy regions (LSC, SSC) and two inverted repeat regions (IRa, IRb) [23]. According to literature records, the ancient cyanophyte endosymbiosis had chloroplasts with a number of functional genes. However, there have been gene loss or transfer events during the evolution of chloroplasts, such as the absence of *ndh* (NADH dehydrogenase) family genes [24]. This was observed in the other species [25,26]. With the rapid advancement of sequencing technology and its decreasing cost, the demand for chloroplast genomes sequencing has been increasing. Some medicinal plants, including *Dipterygium glaucum*, *Cleome chrysantha*, *Bupleurum sikangense*, and *Ephedra equisetina* have been sequenced with massive chloroplast genome data obtained [25,27,28]. Moreover, chloroplast genome sequences have been widely used in phylogenetic and population genetic studies [29–31].

In this study, we sequenced and annotated the chloroplast genomes of *E. przewalskii* and *E. monosperma*. We also carefully compared them with the other published chloroplast genomes from *Ephedra* (*E. intermedia*, *E. equisetina*, *E. foeminea*, and *E. sinica*) to detect the differences in the chloroplast genome. The analysis of chloroplast genome structure, long repeats, short repeats, codon preference, prediction of potential RNA editing sites, and analysis of the adaptive evolution by selective pressure analysis of protein-coding genes contributes to a better understanding of the differences in the chloroplast genome of *Ephedra* species. It provides valuable and abundant data for the phylogenetic analysis and species identification of *Ephedra*.

## 2. Materials and Methods

### 2.1. DNA Extraction and Sequencing

Fresh leaves of *E. przewalskii* and *E. monosperma* were sampled in Mangai (Geographic coordinates: 38°25' N, 90°48' E; Altitude: 2594 m) and Xinghai (Geographic coordinates: 35°21' N, 99°13' E; Altitude: 2594 m), Qinghai Province, P. R. China, respectively (Table S1). The fresh leaves were cleaned with 75% alcohol and ddH<sub>2</sub>O, quickly placed in liquid nitrogen, then transferred to –80 °C for storage after returning to the laboratory. Voucher specimens (*E. przewalskii*: QXA160729005; *E. monosperma*: Chensl-0514) were deposited in the Qinghai-Tibetan Plateau Museum of Biology (HNWP). Total genomic DNA was extracted from fresh leaf tissue by the modified cetyltrimethylammonium bromide (CTAB) method [32]. Qubit Fluorometer (Thermo Fisher, Asheville, NC, USA) was used to estimate DNA concentration. Quality analysis of extracted DNA was evaluated using agarose gel electrophoresis and completed library preparation, following the manufacturer's instructions. *E. przewalskii* and *E. monosperma* were sequenced on Novaseq 6000 platform (Illumina Inc., San Diego, CA, USA) with 150 bp paired-end (PE) sequencing.

### 2.2. Genome Assembly and Annotation

Raw data were filtered using Trimmomatic v. 0.33 [33] and FastQC v. 0.11.8 [34] by discarding low-quality reads, shorter reads, and adapters. High-quality reads (clean reads) were assembled with the default parameters by using SPAdes v3.13.1 [35] and NOVOPlasty

v3.2, with *E. equisetina* (MH161420) as the reference genome [36]. Online software GeSeq [37] annotated the complete chloroplast genome of two *Ephedra* species with reference genomes (*E. monosperma*, Genbank: NC\_054357 and *E. equisetina*, Genbank: MH161420). After manually reviewing, the GenBank files were submitted to GB2sequin to obtain the original sequin files. Sequin software v16.0 was used to check sequin files by adjusting the position of the intron and exon. The circular gene map of the chloroplast genome was drawn by Organellar GenomeDRAW (OGDRAW) [38], and the final cp genomes of *E. przewalskii* and *E. monosperma* were submitted to the GenBank [39]. Available online: <https://www.ncbi.nlm.nih.gov> (accessed on 3 January 2021) (Accession number: *E. przewalskii* MZ567015 and *E. monosperma* OK505605).

### 2.3. Comparative Plastomics in *Ephedra*

We used the online software BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed on 10 January 2021) to determine cp genome sequences with more than 95% coverage and a length was more than 100 kb. Finally, four other *Ephedra* species were selected to conduct comparative genomic studies, including *E. intermedia* (NC\_044772.1), *E. equisetina* (MH161420), *E. foeminea* (NC\_029347), and *E. sinica* (NC\_044773). *E. przewalskii* (MZ567015) and *E. monosperma* (OK505605) plastome sequences were compared with the above-mentioned *Ephedra* cp genomes to visualize their similarities and differences using mVISTA online software with Shuffle-LAGAN mode, with *E. przewalskii* as the reference [40]. IRscope was used to compare LSC (Large Single Copy), IRb (Inverted repeat), SSC (Small Single Copy), and IRa (Inverted repeat) regions of these six complete cp genomes with default parameters, illustrating the contraction and expansion for IR/SC regions [41]. The GC content of the six species was conducted by MEGAX 11.0 [42].

The Microsatellite identification software (MISA) were used for simple sequence repeats (SSRs) analysis, with search parameters: 1 = 10; 2 = 5; 3 = 4; 4 = 3; 5 = 3; 6 = 3 for mono-, di-, tri-, tetra-, penta-, and hexanucleotide sequences, respectively [43]. We used REPuter software to identify forward (F), reverse (R), complementary (C), and palindromic (P) repeat sequences with a minimal length of 30 bp, Hamming distance of 3, and 90% sequence identity [44].

The sequences of six *Ephedra* complete cp genomes were aligned using MAFFT v7 [45]. Then, we used aligned repeats to calculate the nucleotide variability (Pi) using a sliding window analysis in DnaSP v6, with a window length of 600 bp and a step size of 200 bp [46]. We chose three *Ephedra* species (*E. przewalskii*, *E. monosperma*, *E. intermedia*), *Gnetum luofuense*, and *Cycas szechuanensis* to study evolutionary selection pressure. Ka Ks\_ Calculator v2.0 was used to obtain the ratio of non-synonymous to synonymous rates (Ka/Ks) for shared protein-coding genes (PCGs) of these five species, with genetic code table: 11 bacterial and plant plastid code, the way of calculation: NG [47]. Moreover, we used the PREP online tool (<http://prep.unl.edu/>, accessed on 30 June 2021) to predict the RNA editing sites for the PCGs of two *Ephedra* (*E. przewalskii*, *E. monosperma*) species.

### 2.4. Codon Usage

Relative synonymous codon usage (RSCU) of six *Ephedra* species (*E. przewalskii*, *E. monosperma*, *E. intermedia*, *E. equisetina*, *E. foeminea*, and *E. sinica*) for protein-coding genes was calculated by using CodonW v 1.4.2 (<http://codonw.sourceforge.net/>, accessed on 12 October 2021). Moreover, to determine the level of usage bias of synonymous codons for six *Ephedra* species, we calculated the various indices. It included CAI (codon adaptation index), CBI (codon bias index), ENc (Effective number of codons), GC3s (GC content of the synonymous third codons), T3s (Synonymous third codon thymine content), C3s (Synonymous third codon cytosine content), A3s (Synonymous third codon adenosine content), and G3s (Homonymous third codon guanine content).

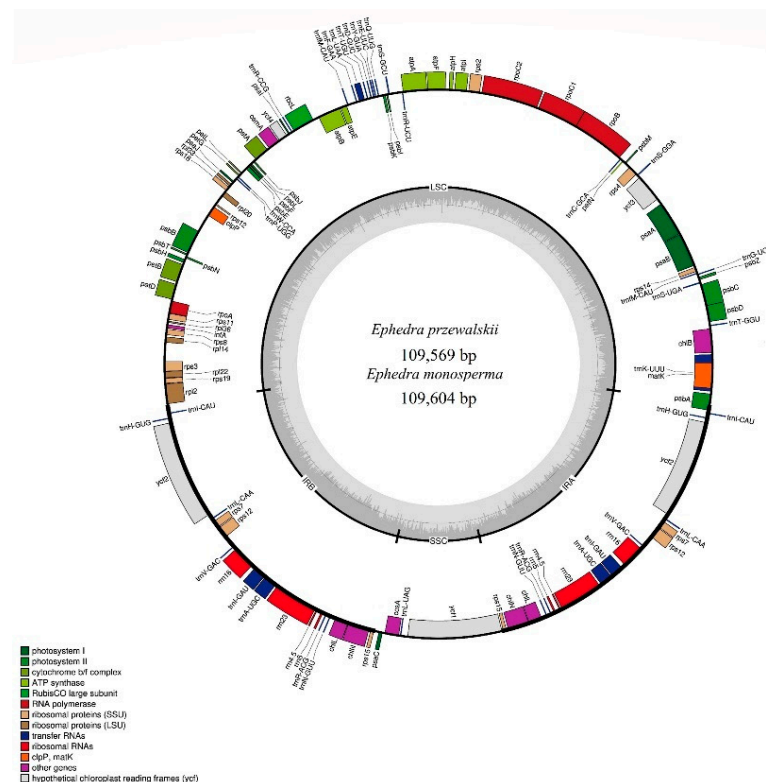
## 2.5. Phylogenetic Profiling

We utilized 14 sequences of *Ephedra* cp genomes from NCBI to conduct phylogenetic analysis on concatenated sequences of 68 PCGs, with *Cycas szechuanensis* (NC\_042668.1) as the outgroup (Table S2). PhyloSuite v1.2.2 was used to extract PCGs of these cp genomes [48]. MAFFT v7 was used to align sequences for CDS (Coding Sequence) and manually adjusted the aligned sequences using MEGA v11.0 [42]. ModelFinder was used to find the best-fitting models in IQ-TREE v1.6.12 [49]. We used IQ-TREE v1.6.12 to reconstruct the Maximum likelihood (ML) tree with the GTR + F + G4 model and MrBayes v3.2.6 in PhyloSuite v1.2.2 to reconstruct the Bayesian inference (BI) tree for the GTR + G + F model, with two parallel runs and 1,000,000 generations [48,50]. Sampling trees every 100 generations, and discarding the first 25% generation (burn-in = 25%) of preheated trees. The branch support analysis was conducted using Ultrafast bootstrap and 5000 bootstrap replications.

## 3. Results

### 3.1. *Ephedra* Chloroplast Genome Features

A total of 10 Gb sequencing data were obtained using Novaseq 6000. All chloroplast genome sequences for *Ephedra* species showed a highly conservative circular structure with four regions, including Large Single Copy (LSC), Small Single Copy (SSC), and two copies of the Inverted Repeat Region (IRa, IRb) (Figure 1). The length of the complete chloroplast genomes in *E. monosperma* (109,604 bp) was longer than *E. przewalskii* (109,569 bp), and the remaining four full-length cp genomes ranged from 109,550 bp in *E. sinica* to 109,667 bp in *E. intermedia* (Table 1). All six *Ephedra* cp genomes were divergent by only 8117 bp in size. The longest length of the LSC region was 60,027 bp in *E. foeminea*, and the shortest was 59,936 bp in *E. intermedia*. The lengths of SSC and IR regions ranged from 8078 bp in *E. equisetina* to 8247 bp in *E. intermedia*, and from 20,731 bp in *E. przewalskii* to 20,753 bp in *E. monosperma*, respectively.



**Figure 1.** Chloroplast genome maps of *E. przewalskii* and *E. monosperma*. The positive coding genes were on the outside of the ring, whereas the reverse coding genes were on the inside of the ring. The darker gray in the inner circle corresponds to GC content, whereas the lighter gray corresponds to AT content. The different groups of genes were shown in different colors.



**Table 1.** Comparison of the chloroplast genome features of six *Ephedra* species.

Species	<i>E. przewalskii</i>	<i>E. monosperma</i>	<i>E. intermedia</i>	<i>E. equisetina</i>	<i>E. foeminea</i>	<i>E. sinica</i>
Accession Number	MZ567015	OK505605	NC_044772.1	MH161420	NC_029347	NC_044773
Genome size (bp)	109,569	109,604	109,667	109,558	109,584	109,550
LSC length (bp)	59,994	60,019	59,936	59,976	60,027	59,961
SSC length (bp)	8113	8079	8247	8078	8079	8103
IR length (bp)	20,731	20,753	20,742	20,752	20,739	20,743
Overall GC content (%)	36.6	36.6	36.6	36.6	36.7	36.7
GC content in LSC (%)	34.2	34.2	34.2	34.2	34.1	34.2
GC content in SSC (%)	27.6	27.9	27.3	27.5	27.7	27.9
GC content in IR (%)	42	42	42.1	42	42	42
Total number of genes	118	118	118	118	118	118
Protein-coding genes	73	73	73	73	73	73
tRNA genes	37	37	37	37	37	37
rRNA genes	8	8	8	8	8	8
Duplicated genes	19	19	18	18	18	18

The overall GC content of six *Ephedra* species varied from 36.6% to 36.7%, in which LSC and SSC regions ranged from 34.1% to 34.2% and from 27.3% to 27.9%, respectively. The IR regions possessed a GC content of 42% in all *Ephedra* species (Table 1). Six *Ephedra* species were identical in gene order and content. A total of 118 genes were annotated, including 73 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Nineteen genes located in IR regions were duplicated, whereas others were unique. The LSC regions contained 58 PCGs and 20 tRNA. The IR regions contained 7 PCGs, 8 tRNA, and four rRNA. The SSC regions had four PCGs, and one tRNA in six *Ephedra* species (Table 2). The *ycf3* gene included two introns. The *ycf3* and *rps12* gene contained three exons, and the remaining ten genes contained two exons. The *rps12* gene was a trans-splicing gene, whose exons were split between LSC and IR regions (Table 3).

**Table 2.** List of annotated genes in six *Ephedra* cp genomes.

Category of Genes	Group of Gene	Gene IDs
Self-replication	Ribosomal RNA genes	<i>rnn23</i> <sup>d,i</sup> ; <i>rnn16</i> <sup>d,i</sup> ; <i>rnn5</i> <sup>d,i</sup> ; <i>rnn4.5</i> <sup>d,i</sup> <i>trnY-GUA</i> <sup>1</sup> ; <i>trnW-CCA</i> <sup>1</sup> ; <i>trnV-GAC</i> <sup>d,i</sup> ; <i>trnT-UGU</i> <sup>1</sup> ; <i>trnT-GGU</i> <sup>1</sup> ; <i>trnS-UGA</i> <sup>1</sup> ; <i>trnS-GGA</i> <sup>1</sup> ; <i>trnS-GCU</i> <sup>1</sup> ; <i>trnR-UCU</i> <sup>1</sup> ; <i>trnR-CCG</i> <sup>1</sup> ; <i>trnR-ACG</i> <sup>d,i</sup> ; <i>trnQ-UUG</i> <sup>1</sup> ; <i>trnP-UGG</i> <sup>1</sup> ; <i>trnN-GUU</i> <sup>d,i</sup> ; <i>trnL-UAA</i> <sup>1</sup> ; <i>trnL-CAA</i> <sup>d,i</sup> ; <i>trnL-AUG</i> <sup>1</sup> ; <i>trnK-UUU</i> <sup>*1</sup> ; <i>trnI-GAU</i> <sup>d,*i</sup> ; <i>trnI-CAU</i> <sup>d,i</sup> ; <i>trnH-GUG</i> <sup>d,i</sup> ; <i>trnG-UCC</i> <sup>1</sup> ; <i>trnM-CAU</i> <sup>d</sup> ; <i>trnF-GAA</i> <sup>1</sup> ; <i>trnE-UUC</i> <sup>1</sup> ; <i>trnD-GUC</i> <sup>1</sup> ; <i>trnC-GCA</i> <sup>1</sup> ; <i>trnA-UGC</i> <sup>d,*i</sup> ; <i>trnL-UAG</i> <sup>s</sup>
	Transfer RNA genes	
Genes for Photosynthesis	Small subunit of ribosome	<i>rps19</i> <sup>1</sup> ; <i>rps18</i> <sup>1</sup> ; <i>rps15</i> <sup>d,i</sup> ; <i>rps14</i> <sup>1</sup> ; <i>rps12</i> <sup>d,**,l&amp;ci</sup> ; <i>rps11</i> <sup>1</sup> ; <i>rps8</i> <sup>1</sup> ; <i>rps7</i> <sup>d,i</sup> ; <i>rps4</i> <sup>1</sup> ; <i>rps3</i> <sup>1</sup> ; <i>rps2</i> <sup>1</sup>
	Large subunit of ribosome	<i>rpl36</i> <sup>1</sup> ; <i>rpl33</i> <sup>1</sup> ; <i>rpl22</i> <sup>1</sup> ; <i>rpl20</i> <sup>1</sup> ; <i>rpl14</i> <sup>1</sup> ; <i>rpl2</i> <sup>*1</sup>
	DNA-dependent RNA polymerase	<i>rpoC2</i> <sup>1</sup> ; <i>rpoC1</i> <sup>*1</sup> ; <i>rpoB</i> <sup>1</sup> ; <i>rpoA</i> <sup>1</sup>
	Subunits of photosystem I	<i>psaA</i> <sup>1</sup> ; <i>psaB</i> <sup>1</sup> ; <i>psaC</i> <sup>s</sup> ; <i>psaI</i> <sup>1</sup> ; <i>psaJ</i> <sup>1</sup>
	Subunits of photosystem II	<i>psbA</i> <sup>*1,1</sup> ; <i>psbB</i> <sup>1</sup> ; <i>psbC</i> <sup>1</sup> ; <i>psbD</i> <sup>1</sup> ; <i>psbE</i> <sup>1</sup> ; <i>psbF</i> <sup>1</sup> ; <i>psbH</i> <sup>1</sup> ; <i>psbI</i> <sup>1</sup> ; <i>psbJ</i> <sup>1</sup> ; <i>psbK</i> <sup>1</sup> ; <i>psbL</i> <sup>1</sup> ; <i>psbM</i> <sup>1</sup> ; <i>psbN</i> <sup>1</sup> ; <i>psbT</i> <sup>1</sup> ; <i>psbZ</i> <sup>1</sup> ; <i>psbN</i> <sup>**1</sup>
	Subunits of Cytochrome b/f complex	<i>petA</i> <sup>1</sup> ; <i>petB</i> <sup>1</sup> ; <i>petD</i> <sup>*1</sup> ; <i>petG</i> <sup>1</sup> ; <i>petL</i> <sup>1</sup> ; <i>petN</i> <sup>1</sup>
	Subunits of ATP synthase	<i>atpA</i> <sup>1</sup> ; <i>atpB</i> <sup>1</sup> ; <i>atpE</i> <sup>1</sup> ; <i>atpF</i> <sup>*1</sup> ; <i>atpH</i> <sup>1</sup> ; <i>atpI</i> <sup>1</sup>
	Large subunit of RUBISCO	<i>rbcL</i> <sup>1</sup>
	Maturase	<i>matK</i> <sup>1</sup>
	Envelope membrane protein	<i>cemA</i> <sup>1</sup>
Other genes	Photochlorophyllide reductase subunit B/L/N	<i>chlB</i> <sup>1</sup> ; <i>chlL</i> <sup>d,i</sup> ; <i>chlN</i> <sup>d,i</sup>
	C-type cytochrome synthesis gene	<i>ccsA</i> <sup>s</sup>
	Protease	<i>clpP</i> <sup>1</sup>
Genes of unknown function	Translational initiation factor	<i>infA</i> <sup>1</sup>
	Conserved open reading frames	<i>ycf1</i> <sup>s</sup> ; <i>ycf2</i> <sup>d,i</sup>
	Assembly/stability of photosystem I	<i>ycf3</i> <sup>**1</sup> ; <i>ycf4</i> <sup>1</sup>

Note: <sup>d</sup> Gene with copies, <sup>\*\*</sup> Gene with one intron/ two introns, <sup>l,s,i</sup> Gene located in LSC/SSC/IR region, <sup>l&ci</sup> Gene was a trans-splicing gene, whose exons could be found in the LSC and IR regions.

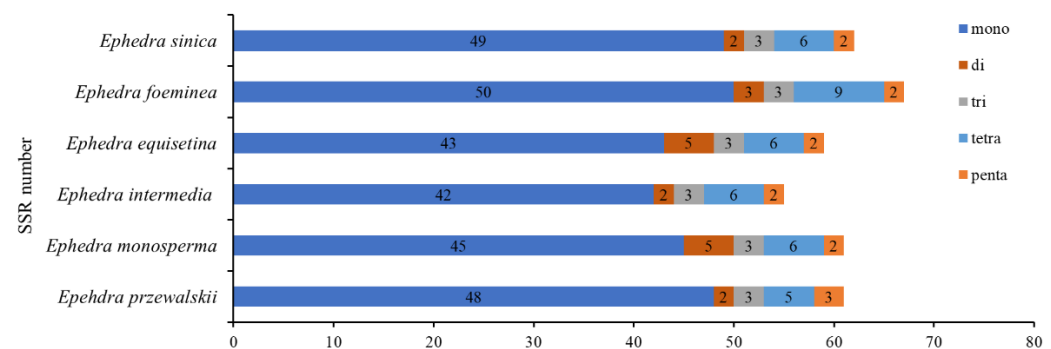
**Table 3.** Exon and intron length of six cp genomes.

Gene	<i>E. przewalskii</i>			<i>E. monosperma</i>			<i>E. intermedia</i>			<i>E. equisetina</i>			<i>E. foeminea</i>			<i>E. sinica</i>		
	Exon I/Intron I	Exon II/Intron II	Exon III	Exon I/Intron I	Exon II/Intron II	Exon III	Exon I/Intron I	Exon II/Intron II	Exon III	Exon I/Intron I	Exon II/Intron II	Exon III	Exon I/Intron I	Exon II/Intron II	Exon III	Exon I/Intron I	Exon II/Intron II	Exon III
<i>rps12</i> *	113/-	31/-	231	113/-	31/-	231	113/-	31/-	231	113/-	31/-	231	113/-	31/-	231	113/-	31/-	231
<i>rpl2</i>	439/499	360/-	-	439/499	360/-	-	439/493	360/-	-	439/499	360/-	-	439/497	360/-	-	439/494	360/-	-
<i>rpl16</i>	395/604	8/-	-	395/608	8-	-	395/603	8/-	-	395/603	8/-	-	395/604	8/-	-	395/602	8/-	-
<i>rpoC1</i>	457/575	1629/-	-	457/581	1629/-	-	457/580	1629/-	-	457/581	1629/-	-	457/581	1629/-	-	457/574	1629/-	-
<i>petB</i>	5/517	641/-	-	5/517	641/-	-	5/517	641/-	-	5/517	641/-	-	5/518	641/-	-	5/517	641/-	-
<i>petD</i>	7/521	474/-	-	7/527	474/-	-	7/525	477/-	-	7/524	477/-	-	7/524	477/-	-	7/518	477/-	-
<i>atpF</i>	143/588	410/-	-	143/584	410/-	-	143/585	410/-	-	143/584	410/-	-	144/590	409/-	-	143/589	410/-	-
<i>ycf3</i>	154/661	225/636	125	154/661	225/615	125	152/658	227/633	125	152/659	227/615	125	152/654	227/636	125	152/655	227/641	125
<i>trnK-UUU</i>	34/2298	37/-	-	34/2298	37/-	-	34/2298	37/-	-	34/2298	37/-	-	34/2294	37/-	-	34/2298	37/-	-
<i>trnI-GAU</i>	35/749	35/-	-	35/761	35/-	-	34/757	36/-	-	34/761	36/-	-	34/755	36/-	-	34/749	36/-	-
<i>trnA-UGC</i>	35/761	39/-	-	35/759	39/-	-	35/758	39/-	-	35/759	39/-	-	35/762	39/-	-	35/760	39/-	-
<i>trnL-UAA</i>	34/291	49/-	-	34/291	49/-	-	34/291	49/-	-	34/291	49/-	-	34/290	49/-	-	34/291	49/-	-

Note: (\*): indicated the *rps12* gene was a trans-splicing gene, and consist of three exons. These exons were split between LSC and IR regions; the number indicates exon and intron length (bp); (-): suggested the absence of intron or gene in the species.

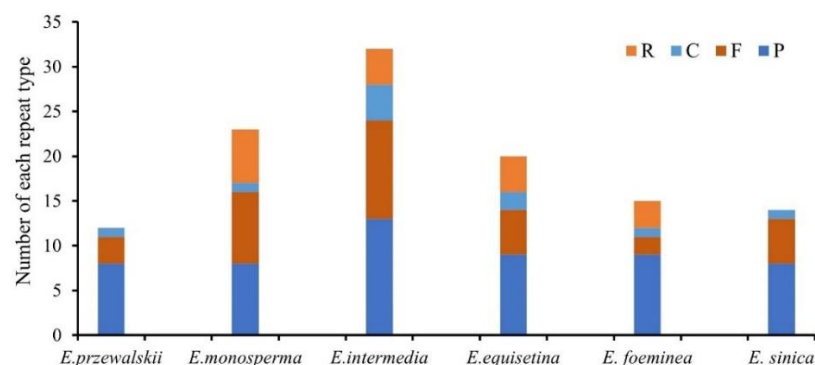
### 3.2. Repeat Sequences and SSR Analysis

We analyzed six cp genome sequences for short repeats (SSRs, simple sequence repeats). The result indicated mononucleotide was the most abundant repeat type, but no hexanucleotide was found in *Ephedra*. 61, 61, 55, 59, 67, and 62 SSRs were detected in *E. przewalskii*, *E. monosperma*, *E. intermedia*, *E. equisetina*, *E. foeminea*, and *E. sinica*, respectively (Figure 2). Among these SSRs, there was the most mononucleotide with the number in *E. foeminea* (50) and the least in *E. intermedia* (42). Besides, we detected SSR in various regions of cp genomes, including LSC, SSC, IR, CDS, rRNA, tRNA, and IGS (Intergenic spacers) (Figure S1). SSRs were more abundant in IGS regions than in other regions, but in rRNA regions, they were the least abundant. The mononucleotide repeat analysis results were presented in six species (*E. przewalskii*, *E. monosperma*, *E. intermedia*, *E. equisetina*, *E. foeminea*, *E. sinica*): the highest number was poly A/T, ranging from 40 to 47, and the lowest was poly C/G, varying from one to three. Only one di-nucleotide (AT/TA) and tri-nucleotide (ATA/TTA), five tetra-nucleotide (AGGT/ATTG, CAAA/TTCT, ATAA/ATCT, ATAG/AATA, and CTAC/CTAT), two pentanucleotide (TTTTA/TTTTTC, ATAAA/AAGAA) were discovered in each cp genomes, while ATTTC was merely in *E. przewalskii* (Figure S2).



**Figure 2.** The type and amounts of simple sequences repeat (SSRs) in the cp genomes of six *Ephedra* species.

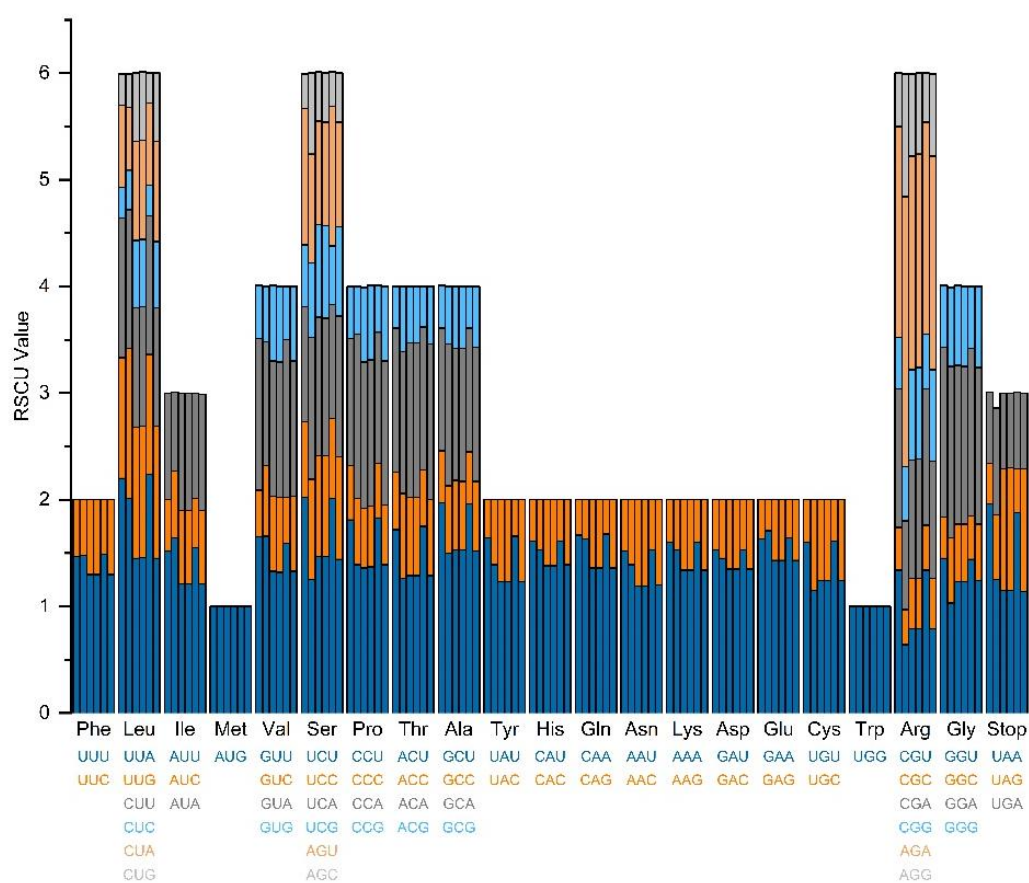
Also, we conducted the long repeats analysis that detected 125 non-overlapped repeats (54 palindromic, 34 forward, 19 complement, and 18 reverse repeats) in six *Ephedra* cp genomes (Figure 3). The palindromic repeats were the most common in these genomes, and the number varies from seven to eight. No reverse repeat was detected in *E. przewalskii*, whereas seven reverse repeats were found in *E. monosperma*. All four types of repeats were more abundant in the LSC region than in the SSC and IR regions. Moreover, we identified the length statistics of long repeat sequences in different size ranges for these cp genomes (Figure S3).



**Figure 3.** The types of long repeats and their number in the cp genome of six *Ephedra* species. R: reverse repeats; C: complement repeats; F: forward repeats; P: palindromic repeats. The number of four types of long repeats were shown in different colors.

### 3.3. Codon Usage

We used 73 shared PCGs to analyze the codon usage bias and to calculate the relative synonymous codon usage (RSCU) value by codonW in all six cp genomes (Figure 4). There were twenty Amino acids and 64 codons, including three stop codons UAA, UAG, and UGA. There was only one codon in Methionine (Met) and Tryptophan (Trp). This study illustrated that there were 22,897 codons in *E. przewalskii*, 27,414 codons in *E. monosperma*, 27,631 codons in *E. intermedia*, 27,625 codons in *E. equisetina*, 27,620 codons in *E. foeminea*, and 27,623 codons in *E. sinica* encoded 73 PCGs, respectively. Moreover, in the six cp genomes, Leucine (Leu) was the most frequent amino acid with the codon number ranging from 2038 to 3603. Trp was the least frequent amino acid with the codon numbers varying from 371 to 490. As in Figure 4, the results of RSCU values indicated slight differences among *E. przewalskii* and *E. monosperma*. The RSCU value of 30 codons was more significant than 1 with A/T endings. Other codons were less than 1 with G/C endings, and the value was equal to 1 with only one codon.



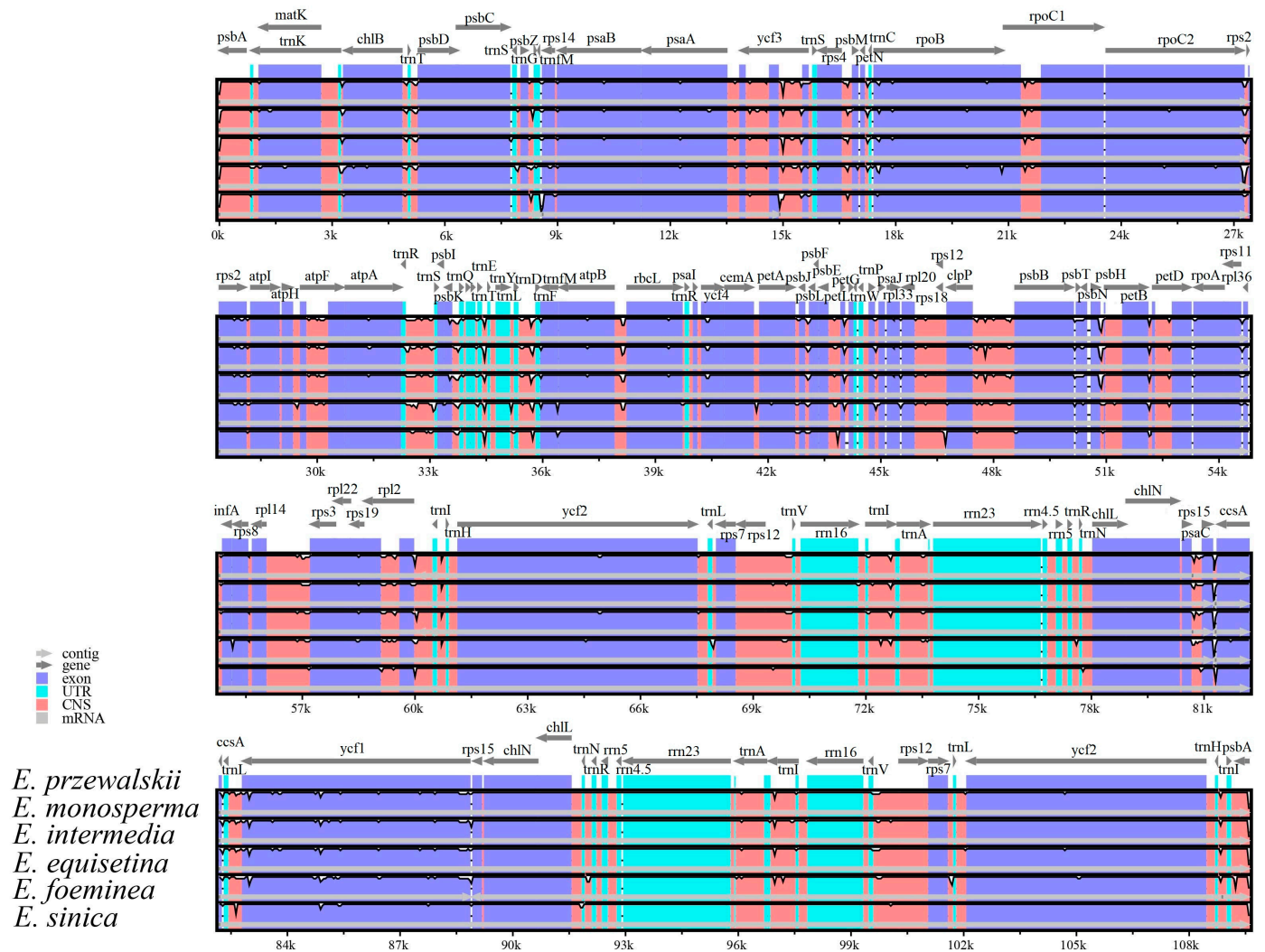
**Figure 4.** Codon usage in 73 protein-coding genes of the cp genomes of six *Ephedra* species. The order of six columns of every amino acid is *E. przewalskii*, *E. monosperma*, *E. intermedia*, *E. equisetina*, *E. foeminea*, and *E. sinica*, respectively.

The results exhibited various indices in the usage bias of synonymous codons for six *Ephedra* species, including CAI, CBI, ENc, GC3s, T3s, C3s, A3s, and G3s (Table S3). Within six *Ephedra* species, the ENc ranged from 46.19 to 54.63, and the GC3s varied from 22.3% to 34.1%. CAI values were 16.6–17.5%.

### 3.4. Divergence in Six *Ephedra* Chloroplast Genome

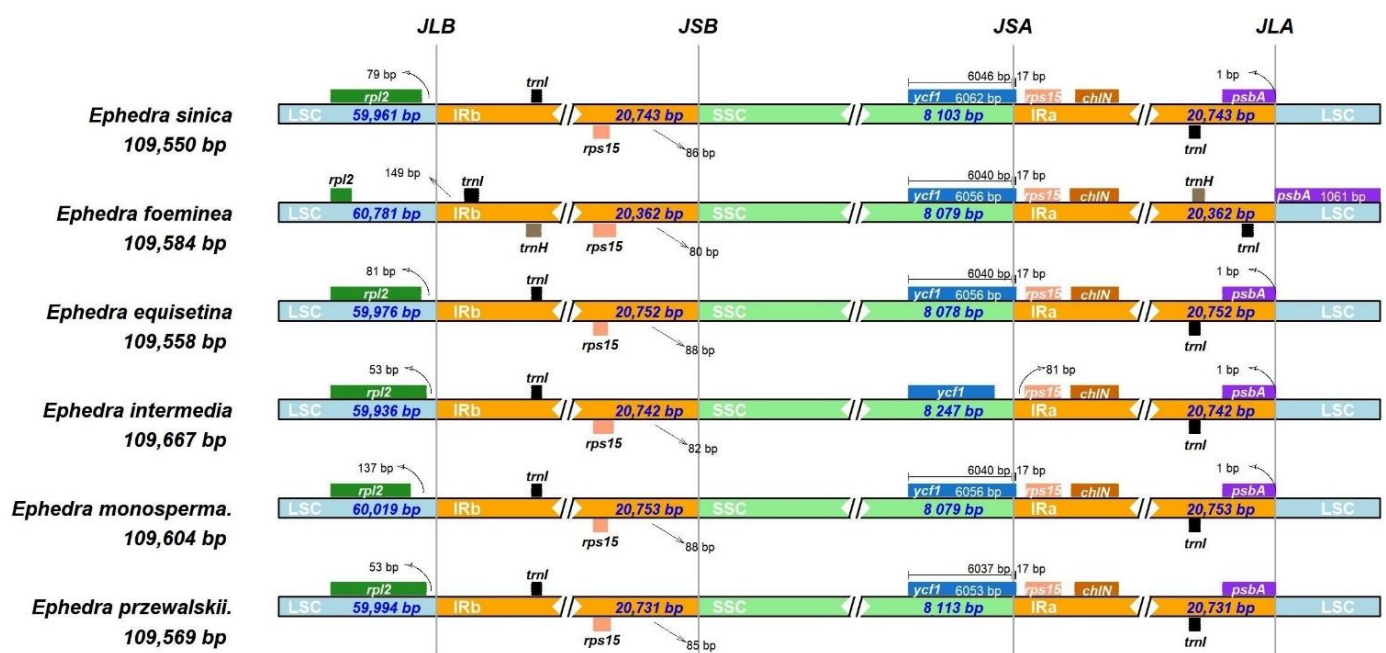
Results of mVISTA revealed, that non-coding regions were less conserved than protein-coding regions (Figure 5). The LSC and SSC regions were more divergent than the IRs regions, and the rRNA gene was highly conserved.





**Figure 5.** Comparison of the chloroplast genome of *E. przewalskii*, *E. monosperma*, *E. intermedia*, *E. equisetina*, *E. foeminea*, *E. sinica*, with *E. przewalskii* as the reference by mVISTA tool.

Also, we compared the boundaries of LSC, SSC, and IR among six cp genomes with IRscope tools. The result exhibited that six *Ephedra* species had little variations of IR/LSC and IR/SSC junction position and characteristics (Figure 6). In five *Ephedra* species, the *rpl2* gene entirely existed in the LSC region. It was 53 to 149 bp away from the LSC/IRb junction regions, except that *E. foeminea* had an extremely short length of the *rpl2* gene. Three genes (*trnI*, *rps15*, and *chlN*) were entirely situated at IR (IRa, IRb) regions, and the *rps15* gene was 80 to 88 bp away from the IRb/SSC junction regions. The *psbA* gene was located in the IRa region of the other *Ephedra* species, but it moved to the LSC region in *E. foeminea*. The *trnH* gene only appeared in the IR region of *E. foeminea*. The *ycf1* gene spanned through the IRa/SSC junction regions and ranged from 6053 to 6062 bp in length. This gene extended by 17 bp of the same length into the IRa regions in five *Ephedra* species. Due to the contraction of the *ycf1* gene in *E. intermedia*, the length was shorter than its whole length (6056 bp)



**Figure 6.** Comparisons of LSC, SSC, and IR region junctions among the six *Ephedra* cp genomes. The numbers above genes represent the distance between the gene and the junction.

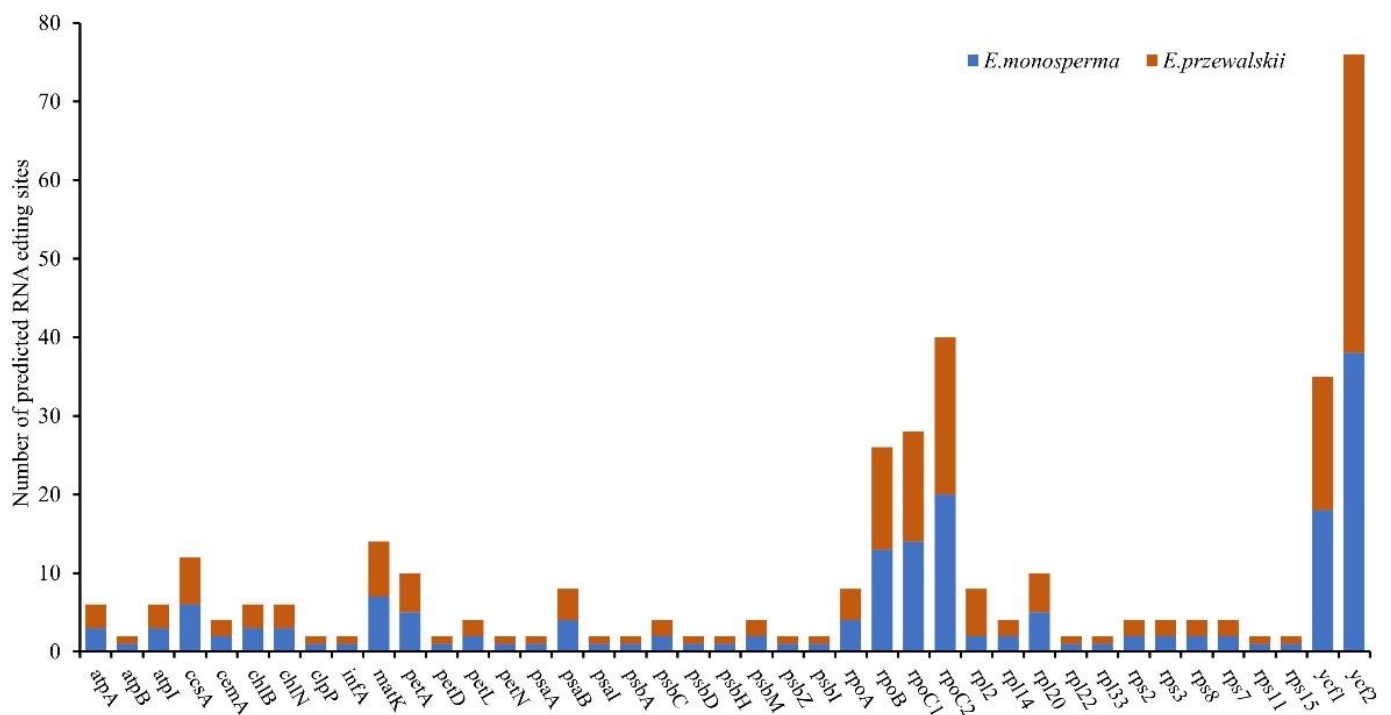
We detected the nucleotide polymorphisms in the complete chloroplast genomes sequence of six *Ephedra* species (Figure S4). The pi value ranged from 0.0000 to 0.018. We found the intergenic spacers with high sequence variability (Pi value > 0.009), including *trnF-GAA\_trnfM-CAU* (0.01389), *trnfM-CAU\_atpE* (0.01189), *rpl36\_infA* (0.11), *psbC\_trnS-UGA* (0.01056), *trnS-UGA\_psbZ* (0.00967), *trnfM-CAU\_rps14* (0.00933). All the above-mentioned intergenic spacers were located in the LSC region. The Pi value (Pi value > 0.009) in the sequence of *ccsA*, *ycf1*, *psaC*, *atpE* and *rpoB* gene was 0.018, 0.01656, 0.01089, 0.01, and 0.00933, respectively. Most of them were found in the SSC region.

### 3.5. Evolutionary Rates in Protein-Coding Genes of *Ephedra* species

We used 68 PCGs from five cp genomes (*E. przewalskii*, *E. monosperma*, *E. intermedia*, *Gnetum luofuense*, *Cycas szechuanensis*), with *Cycas szechuanensis* as the reference. The final results were the average values of non-synonymous nucleotide substitutions (Ka), synonymous nucleotide substitutions (Ks), and Ka/Ks for 68 PCGs, respectively (Table S4). Among the Ks value of genes, 0.0005 in *ycf2* and 0.063 in *psbT* were the smallest and the largest value, respectively. All PCGs showed a lower Ka value. The smallest Ka value was 0.0002 in *psaB*, and the largest was 0.007 in *rps12* and *rps14*. The Ka/Ks values of 68 PCGs were less than 1, meaning that the synonymous substitutions rates were higher than the non-synonymous nucleotide substitutions. It exhibited that all of them undergo purifying selection, ranging from 0 to 0.97. The largest Ka/Ks value was in *rpoA*, and the smallest was in *cemA* and *chlB*.

### 3.6. Predicted RNA Editing Sites for *E. przewalskii* and *E. monosperma*

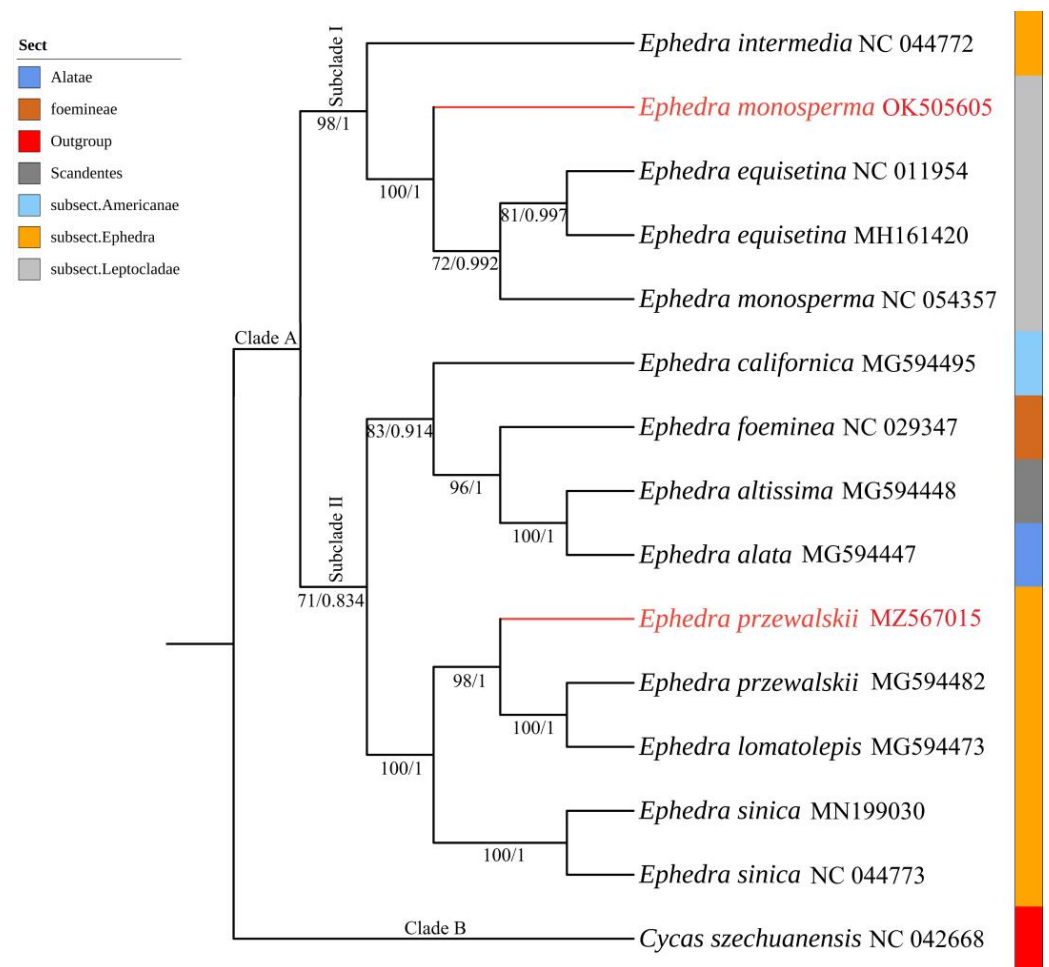
We predicted the RNA editing sites for *E. przewalskii* and *E. monosperma* (Figure 7; Table S5). We found 57, and 56 predicted RNA editing sites in the 15 PCGs of *E. przewalskii* and *E. monosperma*, respectively. Among these PCGs, the RNA Polymerase group had the highest number of predicted RNA editing sites. Specifically, the genes *rpoB*, *rpoC2*, *rpoC1*, and *rpoA* possessed fourteen, eleven, nine, and two RNA editing sites, respectively. All predicted editing sites were C to U transitions, and the most frequent amino acid conversions were from proline to serine.



**Figure 7.** The number of predicted RNA editing sites in the cp genomes of *E. przewalskii* and *E. monosperma*. The X-axis presents the PCGs with predicted editing sites. Y-axis presents the number of predicted editing sites for each PCGs.

### 3.7. Phylogenetic Inference

We inferred the phylogenetic relationship of 14 cp genomes sequences of *Ephedra* and observed the same tree topology in maximum likelihood (ML) and Bayesian inference (BI) analysis (Table S2; Figure 8). The maximum likelihood of bootstrap support (MLBS) and bayesian posterior probability (BPP) were high for each lineage. The first monophyletic Clade (Clade A) included all *Ephedra* species (*E. przewalskii*, *E. monosperma*, *E. intermedia*, *E. equisetina*, *E. foeminea*, *E. sinica*, *E. alata*, *E. altissima*, *E. lomatolepis*, *E. californica*). A sister group to the second Clade (Clade B), included *Cycas szechuanensis* from Cycadaceae. However, *E. przewalskii* (MZ567015) and *E. monosperma* (OK505605) were not gathered in one clade, consistent with the classification system by Flora of China [1]. Specifically, Clade A included subclade I (*E. monosperma*, *E. intermedia*, *E. equisetina*) and subclade II (*E. californica*, *E. foeminea*, *E. altissima*, *E. alata*, *E. przewalskii*, *E. lomatolepis*, *E. sinica*). In subclade I, *E. intermedia* (NC\_044772, subsect. *Ephedra*) was sister to *E. monosperma* (OK505605), two *E. equisetina* (NC\_011954, MH161420), and one *E. monosperma* (NC\_054357) (subsect. *Leptocladae*) (MLBS = 98, BPP = 1). In subclade II, *E. californica* (MG594495, subsect. *Americanae*), *E. foeminea* (NC\_029347, sect. *foemineae*), *E. altissima* (MG594448, sect. *Scandentes*), and *E. alata* (MG594447, sect. *Alatae*) were sister to two *E. przewalskii* (MZ567015, MG594482), *E. lomatolepis* (MG594473), and two *E. sinica* (MN199030, NC\_044773) (subsect. *Ephedra*) (MLBS = 71, BPP = 0.834).



**Figure 8.** Phylogenetic relationship of two *Ephedra* species with related species based on the CDS shared by all cp genomes sequence. *E. przewalskii* (MZ567015) and *E. monosperma* (OK505605) were not gathered in one clade, and they were marked by the red color. Trees obtained with ML (Maximum Likelihood) and BI (Bayesian inference) methods have identical topology, therefore number/number above the branch means ML bootstrap support (MLBS)/Bayesian posterior probability (BPP).

#### 4. Discussion

##### 4.1. Genome Feature in *Ephedra*

With the advancement of sequencing technology and cost reduction, chloroplast genome research has been continuously growing, with an increasing number of chloroplast genomes published in public databases, deepening people's understanding and knowledge of plastid genomes [51]. These abundant genomic data paved the way for plant phylogenetic analysis [52,53]. The chloroplast genome had many advantages, such as maternal inheritance, highly conserved, abundant gene composition, etc. [52]. It has evolved into one of the most effective phylogenetic studies and molecular taxonomy tools [52]. The genomic information was extremely valuable in terms of species origin, evolution, and species relationships, and it has solved numerous phylogenetic problems [54,55]. As a traditional Chinese herbal medicine, *Ephedra* had crucial medicinal value [56]. We revealed the phylogenetic relationship between ten *Ephedra* species based on shared chloroplast PCGs, providing valuable phylogenetic information for this genus. At the same time, conducting comparative genomics and evolutionary analysis can provide abundant data from the plastid genome for species identification of *Ephedra*.

High AT base content found in the cp genomes of six *Ephedra* species was also reported in gymnosperms [25]. The GC content ranged from 36.6% to 36.7%, with that of IR regions significantly higher than that of the SC regions in the six species (Table 1). There were



several examples consistent with previous studies [53,57,58]. Gene loss, duplication, and transfer between chloroplast and nuclear genomes, reliable sources of evolution as they are, also occurred in this study [59]. These events also occurred in this study. It may offer helpful information on evolutionary studies for *Ephedra* and whole gymnosperm plants.

*Ndh* (NADH dehydrogenase subunit) encoded the subunits of the proton-pumping NADH and played a significant role in green plants [60]. The *Ephedra* species in this study had lost all the *ndh* family genes, consistent with the result of the same genus and related genus, *Welwitschia* [25,26]. Also, we believed that there might be two reasons for the loss of *ndh* genes in the *Ephedra* species. First, this type of gene did not play any role in the evolution of *Ephedra* and was eliminated after selection. Secondly, the absence of *ndh* genes in some gymnosperms, including Pinaceae and Gnetales, was related to the living conditions in the Mesozoic era, including high temperature and carbon dioxide concentration [61–63]. However, there was no relevant evidence to prove that the missing *ndh* functional genes might be transferred to the nucleus to become a nuclear gene due to the high mutation rate. The absence of genes did not appear to affect the photosynthetic function of green plants [64,65]. Evidence has shown that *Welwitschia* plants mainly rely on glyceric acid to metabolize CAM for photosynthesis due to the loss event of the *ndh* genes [60]. Therefore, we assumed that *Ephedra* also depends on this metabolism to replace the function of the *ndh* genes. Losing genes can be a standard feature of the *Ephedra* species to clarify its phylogenetic relationship. Further studies with increasing samples are necessary to provide more information for the evolutionary analysis of *Ephedra*. The gene *matK* was one of the fastest evolving genes in the cp genome and located in the intron of *trnK-UUUU* [66]. This gene was broadly used in phylogenetic studies within families, and inter-genera [67–70]. This gene had a 0.751 Ka/Ks value in this study, indicating it experienced purifying selection.

#### 4.2. Comparison of Genomes for *Ephedra*

Six cp genomes were carefully compared to determine their identity and divergence. In mVISTA, the plot could visualize the sequence identity of six *Ephedra* species (Figure 2). The coding region of six cp genomes was more conservative than the non-coding region, congruent with other studies [71]. We found 31 non-coding highly variable regions, and nine PCGs were also highly variable. These highly variable regions and genes might be molecular markers for *Ephedra* species identification and population genetic study. There were subtle divergences in the LSC/IRS/SSC boundary gene distribution in the six *Ephedra* species. The *psbA* gene moved to the LSC region in *E. foeminea*, its length was longer than other species. This gene was a complete gene in *E. foeminea*. Compared with other *Ephedra* species, *E. foeminea*'s IR region and *rpl2* gene in length were shorter, with its IR region showing noticeable contraction changes. The *rpl2* gene was a pseudogene in *E. foeminea*. The *ycf1* gene was entirely located in the SSC region of *E. intermedia*, it was a pseudogene in this species. These results can be used as one of the features to distinguish *Ephedra* species.

Nucleotide diversity was a proposed measure, to express the degree of nucleotide polymorphism in a population [72]. We analyzed the sequence variation of six complete cp genomes. The IR regions were found to be more conserved than the SC regions. The nucleotide polymorphisms in the SC region were greater than that in the IR region. The sequence of the *ccsA* gene with the highest Pi value (Pi = 0.018) was located in the SSC region. The *trnF-GAA\_trnfM-CAU* intergenic spacer showed the highest variation (Pi = 0.01389) and was located in the LSC region.

SSRs are highly variable genetic markers and could be used for species identification, population genetics analyses, or evolutionary biology studies [73,74]. Our SSRs analysis revealed that the single-base repeat type (A/T) variation was the most abundant, implying that there were more replications in six *Ephedra* species. SSRs were mainly found in the LSC and the IGS, while fewer SSRs were situated in the IR region. This result has been observed in other plants [75]. These results may be used to study genetic diversity for *Ephedra* species. The number and distribution of the four types of long repeats differed lightly between



the six *Ephedra* species. All species had complement, forward, and palindromic repeats. The reverse repeat was also identified in all *Ephedra* species except for *E. przewalskii* and *E. sinica*.

#### 4.3. Codon Usage Bias Analysis

Understanding codon usage bias might reveal the effects of long-term evolution on the plant genome [76]. Due to the combined effects of gene selection, mutation, and drift in the long-term evolution process, most species had various Codon Usage Bias [77]. Interestingly, we found that codon usage preference mostly ends with AT, consistent with the determination result of AT-rich base content. CAI could estimate gene expression levels as an essential indicator of species codon usage preference [78,79]. These values were all-around 0.175, indicating a relatively low codon usage preference. The parameter ENc could quantify codon usage bias with a greater than 46 value in *Ephedra* species, indicating that codon preference was weak [80]. Similar results were observed for other species [81]. In general, these indicators analysis found that codon usage preferences of *Ephedra* species were not strong. Still, this study was limited to these few parameters, which were insufficient to explain the preference strength of *Ephedra* species. Therefore, the sample of *Ephedra* should be increased for a more specific analysis of codons usage bias in future studies.

Since the discovery of the first RNA editing event in trypanosome mitochondria [82], many studies on RNA editing have been published [83,84]. RNA editing events also existed in plants, which affect plants' growth, development, and response to various stresses [85–88]. For most of the chloroplast genomes of gymnosperms and angiosperms, only a few dozen RNA editing sites could be predicted [89]. Similarly, we observed only 56–57 RNA editing sites in the cp genomes of *E. przewalskii* and *E. monosperma*. The most frequent type of RNA editing was from C to U conversion in a variety of plants such as thale cress (*Arabidopsis thaliana*) [90], grape (*Vitis vinifera*) [91], tobacco (*Nicotiana tabacum*) [92], as well as this two *Ephedra* species. Our study indicated that the most frequent amino acid conversion was proline to serine in the two *Ephedra* species, similar to the results in other studies [56,59]. Moreover, all the DNA-dependent RNA polymerase genes involved the most predicted RNA editing sites. These results are published for the first time and can offer a novel insight into future RNA editing studies for *E. przewalskii* and *E. monosperma*.

#### 4.4. Evolution Analysis

Analysis of the non-synonymous and synonymous substitutions ratio has become a significant element of molecular evolution studies [93]. It was widely utilized to determine selection pressure for PCGs [94]. The Ka/Ks value of 68 PCGs was less than 1, meaning non-synonymous substitutions were less than synonymous substitutions. This often results in harmful traits for non-synonymous substitutions that would face elimination [95]. These PCGs were subject to purifying selection, similar to that in other studies [27,96,97]. The DNA-dependent RNA polymerase function gene group of *E. przewalskii* and *E. monosperma* showed a high Ka/Ks value, especially the *rpoA* with the highest value (0.97). This illustrated that these genes had a faster evolution rate than other functional groups of genes.

#### 4.5. Phylogenetic Analysis

Many scholars have explored the phylogenetic relationship of *Ephedra*, mainly based on RAPD markers, nrDNA, and cpDNA sequences [11,16,17,98]. This study also utilized the concatenated sequences of chloroplast PCGs to construct phylogenetic trees based on different methods (ML/BI), with high bootstrap value and Bayesian posterior probability. All *Ephedra* species clustered together in a well-supported monophyletic clade, congruing with previous studies [19,20]. In our study, *E. przewalskii* and *E. monosperma* did not cluster together, consistent with the previously reported phylogenetic trees based on molecular data [18,20,99]. Moreover, from the phylogenetic tree's topological structure, it could be observed that *Ephedra* species from different countries are nested together. The *Ephedra*

species (*E. californica*, *E. altissima*, *E. alata*) from North America, Africa, and Kazakhstan were all nested in the sister branches of China species (*E. foeminea*), consistent with previous findings [99]. In summary, we reported the phylogenetic relationship of the ten *Ephedra* species based on the concatenated sequences of PCGs. It offers valuable information for the phylogeny of the genus *Ephedra*.

## 5. Conclusions

Our study reported the complete chloroplast genome of *E. przewalskii* and *E. monosperma*. We also revealed the slight differences in cp genome characterization, the number of long repeats and SSRs, codon usage bias, and RNA editing sites. Moreover, all PCGs in *Ephedra* species were affected by purifying selection. Also, we revealed the phylogenetic relationship of ten *Ephedra* species with high maximum likelihood bootstrap support and Bayesian posterior probability. These data provide valuable and abundant information for the phylogenetic analysis and species identification of the medical plant, *Ephedra*.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/d14100792/s1>. Figure S1: The number of short sequence repeats (SSRs) in LSC, SSC, IR, CDS, rRNA, tRNA, and IGS in the cp genomes of six *Ephedra* species; Figure S2: The frequency of each SSR in the cp genomes of six *Ephedra* species. The order of six columns of each SSR is *E. przewalskii*, *E. monosperma*, *E. intermedia*, *E. equisetina*, *E. foeminea*, *E. sinica*, respectively; Figure S3: The number of four types of long repeats with different lengths ranges in the cp genomes of six *Ephedra* species; R: reverse repeats; C: complement repeats; F: forward repeats; P: palindromic repeats. Repeats with different lengths for six cp genomes are shown in different colors; Figure S4: The results of the nucleotide variability (Pi) values in the six *Ephedra* species. The x-axis is the position of the midpoint of the window. The y-axis is the nucleotide diversity (Pi) values of each window; Table S1: Specimen information of *E. przewalskii* and *E. monosperma*; Table S2: The cp genome sequences used in the phylogenetic analysis; Table S3: Codon usage of protein-coding genes in chloroplast genomes of six *Ephedra* species (Unit: %); Table S4: The Ka/Ks ratio of 68 protein-coding genes of five cp genomes (*E. przewalskii*, *E. monosperma*, *E. intermedia*, *Gnetum luofuense*, *Cycas szechuanensis*), with *C. szechuanensis* as the reference; Table S5: The number of predicted RNA editing sites in the cp genomes of *E. przewalskii* and *E. monosperma*.

**Author Contributions:** Conceptualization, S.H., M.X. and J.Y.; methodology, S.H. and M.X.; software S.H. and J.Y.; validation, S.H., J.Y. and F.Z.; formal analysis, S.H.; investigation, M.X., J.Y., H.X. and F.Z.; resources, M.X., J.Y., H.X. and F.Z.; data curation, S.H., H.X. and Y.H.; writing—original draft preparation S.H. and F.Z.; writing—review and editing, all authors; visualization, S.H. and H.X.; supervision, F.Z.; project administration, F.Z.; funding acquisition, F.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Second Tibetan Plateau Scientific Expedition and Research (STEP) Program (2019QZKK0502), Construction Project for Innovation Platform of Qinghai Province (2022-ZJ-Y04).

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** All data generated or analyzed during this study are included in this published article.

**Acknowledgments:** The authors are most grateful to anonymous reviewers for their helpful suggestions and comments.

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

## References

1. Cheng, W.J.; Fu, L.K. *Gymnospermae. Flora of China*; Science Press: Beijing, China, 1978; Volume 7, pp. 471–489.
2. Qin, A.L.; Wang, M.M.; Cun, Y.Z.; Yang, F.S.; Wang, S.S.; Ran, J.H.; Wang, X.Q. Phylogeographic evidence for a link of species divergence of *Ephedra* in the Qinghai-Tibetan Plateau and adjacent regions to the Miocene Asian aridification. *PLoS ONE* **2013**, *8*, e56243. [[CrossRef](#)] [[PubMed](#)]

3. Meyer, C.A. Versuch einer Monographie der Gattung *Ephedra*. *Mem. Akad. Imper. Sci. St. Petersburg. Ser.6 (Sci. Nat.)* **1846**, *5*, 225–297. (In German)
4. Stapf, O. Die Arten der Gattung *Ephedra*. *Denkschr. Math.-Nat. Kl. Akad. Wiss. Wien, lvi.* **1889**, *56*, 35.
5. Shen, G.M. Distribution and evolution of the genus *Ephedra* in China. *Arid. Zone Res.* **1993**, *10*, 39–48.
6. Yang, Y. Systematics and Evolution of *Ephedra* L. (Ephedraceae) from China. *Chin. Acad. Sci.* **2002**, 1–231. Available online: <https://cir.nii.ac.jp/crid/1572824500456924160> (accessed on 3 January 2021).
7. Vishal, S.; Harihara, V.; Mehendale, M. *Ephedra*. In *Encyclopedia of Toxicology*, 2nd ed.; Philip, W., Ed.; Elsevier: Amsterdam, The Netherlands, 2005; pp. 223–228.
8. Barker, W.D.; Antia, U. A study of the use of *Ephedra* in the manufacture of methamphetamine. *Forensic Sci. Int.* **2007**, *166*, 102–109. [[CrossRef](#)]
9. Haller, C.A.; Benowitz, N.L.; Jacob, P. Hemodynamic effects of ephedra-free weight-loss supplements in humans. *Am. J. Med.* **2005**, *118*, 998–1003. [[CrossRef](#)]
10. Mei, J.; Zhou, Y.; Yang, X.; Zhang, F.; Liu, X.; Yu, B. Active components in *Ephedra sinica* Stapf disrupt the interaction between ACE2 and SARS-CoV-2 RBD: Potent COVID-19 therapeutic agents. *J. Ethnopharmacol.* **2021**, *278*, 114303. [[CrossRef](#)]
11. Jiang, H.Y.; Li, W.; Li, S.; Shi, L.N.; Gao, Y.H.; Wu, J.L. Relationship analysis of five species in the genus *Ephedra* L. by RAPD. *J. Gansu Agric. Univ.* **2006**, *41*, 49–52.
12. Deng, N.; Shi, S.Q.; Chang, E.M.; Liu, J.F.; Lan, Q.; Jiang, Z.P. Transcriptomic Analysis of Germinated Seeds of *Ephedra przewalskii*. *J. Northeast. For. Univ.* **2015**, *43*, 28–32.
13. Xin, G.-Z.; Hu, B.; Shi, Z.-Q.; Zheng, J.-Y.; Wang, L.; Chang, W.-Q.; Li, P.; Yao, Z.-P.; Liu, L.-F. A direct ionization mass spectrometry-based approach for differentiation of medicinal *Ephedra* species. *J. Pharm. Biomed. Anal.* **2016**, *117*, 492–498. [[CrossRef](#)] [[PubMed](#)]
14. Zheng, Y.; Gao, H.; Song, M.; Lin, Y.; Fan, J.; Liu, X. Identification of plant materials containing ephedrine alkaloids based on DNA barcoding and TaqMan real-time PCR assay. *Acta Physiol. Plant.* **2021**, *4311*, 143. [[CrossRef](#)]
15. Rydin, C.; Pedersen, K.R.; Friis, E.M. On the evolutionary history of *Ephedra*: Cretaceous fossils and extant molecules. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16571–16576. [[CrossRef](#)] [[PubMed](#)]
16. Long, C.; Kakiuchi, N.; Takahashi, A.; Komatsu, K.; Cai, S.; Mikage, M. Phylogenetic analysis of the DNA sequence of the non-coding region of nuclear ribosomal DNA and chloroplast of *Ephedra* plants in China. *Planta Med.* **2004**, *70*, 1080–1084. [[CrossRef](#)] [[PubMed](#)]
17. Huang, J.; Giannasi, D.E.; Price, R.A. Phylogenetic relationships in *Ephedra* (Ephedraceae) inferred from chloroplast and nuclear DNA sequences. *Mol. Phylogenet. Evol.* **2005**, *35*, 48–59. [[CrossRef](#)] [[PubMed](#)]
18. Chen, X.; Cui, Y.; Nie, L.; Hu, H.; Xu, Z.; Sun, W.; Gao, T.; Song, J.; Yao, H. Identification and Phylogenetic Analysis of the Complete Chloroplast Genomes of Three *Ephedra* Herbs Containing Ephedrine. *Biomed. Res. Int.* **2019**, *2019*, 5921725. [[CrossRef](#)] [[PubMed](#)]
19. Rydin, C.; Blokzijl, R.; Thureborn, O.; Wikström, N. Node ages, relationships, and phylogenomic incongruence in an ancient gymnosperm lineage—Phylogeny of *Ephedra* revisited. *Taxon* **2021**, *70*, 701–719. [[CrossRef](#)]
20. Moore, M.J.; Soltis, P.S.; Bell, C.D.; Burleigh, J.G.; Soltis, D.E. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4623–4628. [[CrossRef](#)] [[PubMed](#)]
21. Samigullin, T.H.; Logacheva, M.D.; Terenteva, E.I.; Degtjareva, G.V.; Vallejo-Roman, C.M. The plastid genome of *Seseli montanum*: Complete sequence and comparison with plastomes of other members of the *Apiaceae* family. *Biochemistry* **2016**, *81*, 981–985. [[CrossRef](#)]
22. Androsiuk, P.; Jastrzębski, J.P.; Paukszto, Ł.; Makowczenko, K.; Okorski, A.; Pszczółkowska, A.; Chwedorzewska, K.J.; Górecki, R.; Gielwanowska, I. Evolutionary dynamics of the chloroplast genome sequences of six *Colobanthus* species. *Sci. Rep.* **2020**, *10*, 11522. [[CrossRef](#)] [[PubMed](#)]
23. Ravi, V.; Khurana, J.P.; Tyagi, A.K.; Khurana, P. An update on chloroplast genomes. *Syst. Evol.* **2008**, *271*, 101–122. [[CrossRef](#)]
24. Zhang, X.; Sun, Y.; Landis, J.B.; Lv, Z.; Shen, J.; Zhang, H.; Lin, N.; Li, L.; Sun, J.; Deng, T.; et al. Plastome phylogenomic study of Gentianeae (Gentianaceae): Widespread gene tree discordance and its association with evolutionary rate heterogeneity of plastid genes. *BMC Plant Biol.* **2020**, *20*, 340. [[CrossRef](#)] [[PubMed](#)]
25. Wu, C.S.; Lai, Y.T.; Lin, C.P.; Wang, Y.N.; Chaw, S.M. Evolution of reduced and compact chloroplast genomes (cpDNAs) in gnetophytes: Selection toward a lower-cost strategy. *Mol. Phylogenet. Evol.* **2009**, *52*, 115–124. [[CrossRef](#)] [[PubMed](#)]
26. McCoy, S.R.; Kuehl, J.V.; Boore, J.L.; Raubeson, L.A. The complete plastid genome sequence of *Welwitschia mirabilis*: An unusually compact plastome with accelerated divergence rates. *BMC Evol. Biol.* **2008**, *8*, 130. [[CrossRef](#)] [[PubMed](#)]
27. Alzahrani, D.; Albokhari, E.; Yaradua, S.; Abba, A. Complete chloroplast genome sequences of *Dipterygium glaucum* and *Cleome chrysantha* and other Cleomaceae Species, comparative analysis and phylogenetic relationships. *Saudi J. Biol. Sci.* **2021**, *28*, 2476–2490. [[CrossRef](#)] [[PubMed](#)]
28. Xie, X.; Huang, R.; Li, F.; Tian, E.; Li, C.; Chao, Z. Phylogenetic position of *Bupleurum sikangense* inferred from the complete chloroplast genome sequence. *Gene* **2021**, *798*, 145801. [[CrossRef](#)] [[PubMed](#)]
29. Thomson, A.M.; Vargas, O.M.; Dick, C.W. Complete plastome sequence from *Bertholletia excelsa* and 23 related species yield informative markers for Lecythidaceae. *Appl. Plant Sci.* **2018**, *6*, e01151. [[CrossRef](#)]
30. Sun, C.; Chen, F.; Teng, N.; Xu, Y.; Dai, Z. Comparative analysis of the complete chloroplast genome of seven *Nymphaea* species. *Aquat. Bot.* **2021**, *170*, 103353. [[CrossRef](#)]

31. Khan, G.; Zhang, F.; Gao, Q.; Fu, P.C.; Xing, R.; Wang, J.; Liu, H.; Chen, S. Molecular phylogeography and intraspecific divergence of *Spiraea alpina* (Rosaceae) distributed in the Qinghai-Tibetan Plateau and adjacent regions inferred from nrDNA. *Biochem. Syst. Ecol.* **2014**, *57*, 278–286. [[CrossRef](#)]
32. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
33. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
34. Simon, A. *FastQC: A Quality Control Tool for High Throughput Sequence Data*; Babraham Bioinformatics, Babraham Institute: Cambridge, UK, 2010.
35. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [[CrossRef](#)]
36. Nicolas, D.; Patrick, M.; Guillaume, S. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **2016**, *4*, e18.
37. Tillich, M.; Lehwick, 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](#)]
38. 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](#)] [[PubMed](#)]
39. Benson, D.A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2013**, *41*, D36–D42. [[CrossRef](#)] [[PubMed](#)]
40. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, W273–W279. [[CrossRef](#)] [[PubMed](#)]
41. Ali, A.; Jaakko, H.; Peter, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031.
42. Kumar, S.; Stecher, G.; Li, M.; Niyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
43. Sebastian, B.; Thomas, T.; Thomas, M.; Uwe, S.; Martin, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585.
44. 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](#)]
45. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2019**, *20*, 1160–1166. [[CrossRef](#)] [[PubMed](#)]
46. 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](#)] [[PubMed](#)]
47. Wang, D.P.; Zhang, Y.B.; Zhang, Z.; Zhu, J.; Yu, J. KaKs\_Calculator 2.0: A Toolkit Incorporating Gamma-Series Methods and Sliding Window Strategies. *Genom. Proteom. Bioinform.* **2010**, *8*, 77–80. [[CrossRef](#)]
48. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)]
49. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)]
50. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572–1574. [[CrossRef](#)]
51. Armbruster, U.; Pesaresi, P.; Pribil, M.; Hertle, A.; Leister, D. Update on chloroplast research: New tools, new topics, and new trends. *Mol. Plant.* **2011**, *4*, 1–16. [[CrossRef](#)]
52. Khan, A.; Khan, I.A.; Asif, H.; Azim, M.K. Current trends in chloroplast genome research. *Afr. J. Biotechnol.* **2010**, *9*, 3494–3500.
53. Huang, Q.; Liu, Z.X.; Wang, C.; Jing, M.Y.; Liu, J.Q.; Zhou, W.; Kai, G.Y. The Complete Chloroplast Genome Sequences of *Anisodus Acutangulus* and a Comparison with Other Solanaceae Species. *Clin. Complement. Med. Pharmacol.* **2021**, *1*, 100002. [[CrossRef](#)]
54. Zhang, Y.; Yu, J.; Xia, M.; Chi, X.; Khan, G.; Chen, S.; Zhang, F. Plastome Sequencing Reveals Phylogenetic Relationships among *Comastoma* and Related Taxa (Gentianaceae) from the Qinghai-Tibetan Plateau. *Ecol. Evol.* **2021**, *11*, 16034–16046. [[CrossRef](#)] [[PubMed](#)]
55. Huang, Y.; Wang, J.; Yang, Y.; Fan, C.; Chen, J. Phylogenomic Analysis and Dynamic Evolution of Chloroplast Genomes in Salicaceae. *Front. Plant Sci.* **2017**, *8*, 1050. [[CrossRef](#)] [[PubMed](#)]
56. González-Juárez, D.E.; Escobedo-Moratilla, A.; Flores, J.; Hidalgo-Figueroa, S.; Martínez-Tagüena, N.; Morales-Jiménez, J.; Muñoz-Ramírez, A.; Pastor-Palacios, G.; Pérez-Miranda, S.; Ramírez-Hernández, A.; et al. A Review of the *Ephedra* genus: Distribution, Ecology, Ethnobotany, Phytochemistry and Pharmacological Properties. *Molecules* **2020**, *25*, 3283. [[CrossRef](#)] [[PubMed](#)]
57. Shahzadi, I.; Abdullah, M.F.; Ali, Z.; Ahmed, I.; Mirza, B. Chloroplast genome sequences of *Artemisia maritima* and *Artemisia absinthium*: Comparative analyses, mutational hotspots in genus *Artemisia* and phylogeny in family Asteraceae. *Genomics* **2020**, *112*, 1454–1463. [[CrossRef](#)] [[PubMed](#)]



58. Wei, F.; Tang, D.; Wei, K.; Qin, F.; Li, L.; Lin, Y.; Zhu, Y.; Khan, A.; Kashif, M.H.; Miao, J. The complete chloroplast genome sequence of the medicinal plant *Sophora Tonkinensis*. *Sci. Rep.* **2020**, *10*, 12473. [\[CrossRef\]](#)
59. Xiong, A.-S.; Peng, R.-H.; Zhuang, J.; Gao, F.; Zhu, B.; Fu, X.-Y.; Xue, Y.; Jin, X.-F.; Tian, Y.-S.; Zhao, W.; et al. Gene duplication, transfer, and evolution in the chloroplast genome. *Biotechnol. Adv.* **2009**, *27*, 340–347. [\[CrossRef\]](#)
60. Shikanai, T. Chloroplast NDH: A different enzyme with a structure similar to that of respiratory NADH dehydrogenase. *Biochim. Biophys. Acta.* **2016**, *1857*, 1015–1022. [\[CrossRef\]](#)
61. Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M. Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9794–9798. [\[CrossRef\]](#)
62. Ranade, S.S.; García-Gil, M.R.; Rosselló, J.A. Non-functional plastid *ndh* gene fragments are present in the nuclear genome of Norway spruce (*Picea abies* L. Karsch): Insights from in silico analysis of nuclear and organellar genomes. *Mol. Genet. Genomics.* **2016**, *291*, 935–941. [\[CrossRef\]](#)
63. Braukmann, T.W.A.; Kuzmina, M.; Stefanović, S. Loss of all plastid *ndh* genes in Gnetales and conifers: Extent and evolutionary significance for the seed plant phylogeny. *Curr. Genet.* **2009**, *55*, 323–337. [\[CrossRef\]](#)
64. Ruhlman, T.A.; Chang, W.J.; Chen, J.J.; Huang, Y.T.; Chan, M.T.; Zhang, J.; Liao, D.C.; Blazier, J.C.; Jin, X.; Shih, M.C.; et al. NDH expression marks major transitions in plant evolution and reveals coordinate intracellular gene loss. *BMC Plant Biol.* **2015**, *15*, 100. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Lin, C.; Chen, J.J.; Chiu, C.; Hsiao, H.C.; Yang, C.; Jin, X.; Leebens-Mack, J.; Depamphilis, C.W.; Huang, Y.; Chang, W.; et al. Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. *Plant J.* **2017**, *90*, 994–1006. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Hilu, K.W.; Liang, H.P. The *matK* gene: Sequence variation and application in plant systematics. *Am. J. Bot.* **1997**, *84*, 830–839. [\[CrossRef\]](#)
67. Plunkett, G.M.; Soltis, D.E.; Soltis, P.S. Clarification of the relationship between Apiaceae and Araliaceae based on *matK* and *rbcL* sequence data. *Am. J. Bot.* **1997**, *84*, 565–580. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Cbol Plant Working Group. A DNA barcode for land plants. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12794–12797. [\[CrossRef\]](#)
69. Steven, G.N.; Subramanyam, R. Testing plant barcoding in a sister species complex of *pantropical Acacia* (Mimosoideae, Fabaceae). *Mol. Ecol. Resour.* **2009**, *9* (Suppl. S1), 172–180. [\[CrossRef\]](#)
70. Wang, T.; Liu, J.; Guo, Y.; Yuan, N. Phylogenetic relationship among local legumes in Jiangsu Province based on analyses of *matK* gene and ITS sequence. *J. Nanjing Agric. Univ.* **2017**, *40*, 795–803.
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. [\[CrossRef\]](#)
72. Nei, M.; Li, W.H. Mathematical Model for Studying Genetic Variation in Terms of Restriction Endonucleases. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 5269–5273. [\[CrossRef\]](#)
73. Tuler, A.C.; Carrijo, T.T.; Nória, L.R.; Ferreira, A.; Peixoto, A.L.; Ferreira, M.F.D.S. SSR markers: A tool for species identification in *Psidium* (Myrtaceae). *Mol. Biol. Rep.* **2015**, *42*, 1501–1513. [\[CrossRef\]](#)
74. Mohamed, A. El-Esawi. SSR analysis of genetic diversity and structure of the germplasm of faba bean (*Vicia faba* L.). *Comptes Rendus Biol.* **2017**, *340*, 474–480.
75. Bi, Y.; Zhang, M.F.; Xue, J.; Dong, R.; Du, Y.P.; Zhang, X.H. Chloroplast genomic resources for phylogeny and DNA barcoding: A case study on *Fritillaria*. *Sci. Rep.* **2018**, *8*, 1184. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Duan, H.; Zhang, Q.; Wang, C.; Li, F.; Tian, F.; Lu, Y.; Hu, Y.; Yang, H.; Cui, G. Analysis of codon usage patterns of the chloroplast genome in *Delphinium grandiflorum* L. reveals a preference for AT-ending codons as a result of major selection constraints. *Peer J.* **2021**, *9*, e10787. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Ma, L.N.; Cui, P.; Zhu, J.; Zhang, Z.H.; Zhang, Z. Translational selection in human: More pronounced in housekeeping genes. *Biol Direct* **2014**, *9*, 17. [\[CrossRef\]](#)
78. Sharp, P.M.; Li, W.H. The codon adaptation index—A measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **1987**, *15*, 1281–1295. [\[CrossRef\]](#)
79. Supriyo, C.; Prosenjit, P.; Mazumder, T.H. Codon Usage Bias Prefers AT Bases in Coding Sequences Among the Essential Genes of *Haemophilus influenzae*. *Not. Sci. Biol.* **2014**, *6*, 417–421. [\[CrossRef\]](#)
80. Wright, F. The 'effective number of codons' used in a gene. *Gene* **1990**, *87*, 23–29. [\[CrossRef\]](#)
81. Jiang, S.Z.; Lian, H.; Xiong, Y.F.; Zhang, S.; Chen, S.P. Analysis of Codon Bias in Chloroplast Genome of *Castanopsis carlesii*. *Mol. Plant Breed.* **2021**, *26*, 1–12.
82. Benne, R. RNA-editing in trypanosome mitochondria. *Biochim. Biophys. Acta.* **1989**, *1007*, 131–139. [\[CrossRef\]](#)
83. Bock, R.; Kössel, H.; Maliga, P. Introduction of a heterologous editing site into the tobacco plastid genome: The lack of RNA editing leads to a mutant phenotype. *EMBO J.* **1994**, *13*, 4623–4628. [\[CrossRef\]](#)
84. Liu, H.; Wang, Q.; He, Y.; Chen, L.; Hao, C.; Jiang, C.; Li, Y.; Dai, Y.; Kang, Z.; Xu, J.R. Genome-wide A-to-I RNA editing in fungi independent of ADAR enzymes. *Genome Res.* **2016**, *26*, 499–509. [\[CrossRef\]](#)
85. Gualberto, J.M.; LaMattina, L.; Bonnard, G.; Weil, J.H.; Grienenberger, J.-M. RNA editing in wheat mitochondria results in the conservation of protein sequences. *Nature* **1989**, *341*, 660–662. [\[CrossRef\]](#)



86. Wakasugi, T.; Hirose, T.; Horihata, M.; Tsudzuki, T.; Kössel, H.; Sugiura, M. Creation of a novel protein-coding region at the RNA level in black pine chloroplasts: The pattern of RNA editing in the gymnosperm chloroplast is different from that in angiosperms. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8766–8770. [[CrossRef](#)]
87. Wang, W.; Yu, H.; Wang, J.; Lei, W.; Gao, J.; Qiu, X.; Wang, J. The Complete Chloroplast Genome Sequences of the Medicinal Plant *Forsythia suspensa* (Oleaceae). *Int. J. Mol. Sci.* **2017**, *18*, 2288. [[CrossRef](#)] [[PubMed](#)]
88. Kumbhar, F.; Nie, X.; Xing, G.; Zhao, X.; Lin, Y.; Wang, S.; Weining, S. Identification and characterisation of RNA editing sites in chloroplast transcripts of einkorn wheat (*Triticum monococcum*). *Ann. Appl. Biol.* **2018**, *172*, 197–207. [[CrossRef](#)]
89. Mower, J.P. The PREP suite: Predictive RNA editors for plant mitochondrial genes, chloroplast genes and user-defined alignments. *Nucleic Acids Res.* **2009**, *37*, W253–W259. [[CrossRef](#)] [[PubMed](#)]
90. Giegé, P.; Brennicke, A. RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 15324–15329. [[CrossRef](#)] [[PubMed](#)]
91. Picardi, E.; Horner, D.S.; Chiara, M.; Schiavon, R.; Valle, G.; Pesole, G. Large-scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing. *Nucleic Acids Res.* **2010**, *38*, 4755–4767. [[CrossRef](#)] [[PubMed](#)]
92. Grimes, B.T.; Sisay, A.K.; Carroll, H.D.; Cahoon, A.B. Deep sequencing of the tobacco mitochondrial transcriptome reveals expressed ORFs and numerous editing sites out-side coding regions. *BMC Geno.* **2014**, *15*, 31.
93. Nei, M.; Kumar, S. *Molecular Evolution and Phylogenetics*; Oxford University Press: New York, NY, USA, 2000.
94. Yin, K.; Zhang, Y.; Li, Y.; Du, F. Different natural selection pressures on the *atpF* gene in evergreen sclerophyllous and deciduous oak species: Evidence from comparative analysis of the complete chloroplast genome of *Quercus aquifolioides* with other oak species. *Int. J. Mol. Sci.* **2018**, *19*, 1042. [[CrossRef](#)] [[PubMed](#)]
95. Yang, Z.H. *Computational Molecular Evolution*; Oxford University Press: Oxford, UK, 2006; p. 284.
96. Azarin, K.; Usatov, A.; Makarenko, M.; Khachumov, V.; Gavrilova, V. Comparative analysis of chloroplast genomes of seven perennial *Helianthus* species. *Gene* **2021**, *774*, 145418. [[CrossRef](#)] [[PubMed](#)]
97. Huang, R.; Xie, X.; Li, F.; Tian, E.; Chao, Z. Chloroplast genomes of two Mediterranean *Bupleurum* species and the phylogenetic relationship inferred from combined analysis with East Asian species. *Planta* **2021**, *253*, 81. [[CrossRef](#)] [[PubMed](#)]
98. Rydin, C.; Pedersen, K.R.; Crane, P.R.; Friis, E.M. Former diversity of *Ephedra* (Gnetales): Evidence from Early Cretaceous seeds from Portugal and North America. *Ann. Bot.* **2006**, *98*, 123–140. [[CrossRef](#)] [[PubMed](#)]
99. Rydin, C.; Korall, P. Evolutionary Relationships in *Ephedra* (Gnetales), with Implications for Seed Plant Phylogeny. *Int. J. Plant Sci.* **2009**, *170*, 1031–1043. [[CrossRef](#)]