agronomy

MDPI

Article

Complete Plastomes of Ten *Rorippa* Species (Brassicaceae): Comparative Analysis and Phylogenetic Relationships

Ting Ren, Lulu Xun, Yun Jia and Bin Li *

Xi'an Botanical Garden of Shaanxi Province (Institute of Botany of Shaanxi Province), Xi'an 710061, China; renting@xab.ac.cn (T.R.); xunlulu@xab.ac.cn (L.X.); jiayun@xab.ac.cn (Y.J.)
* Correspondence: lb@xab.ac.cn

Abstract: The genus *Rorippa* belongs to the family Brassicaceae, and its members usually have high medicinal value. The genus consists of approximately 75 species and mainly grows in the Northern Hemisphere, occurring in every continent except Antarctica. The taxonomy and phylogenetic relationships of *Rorippa* are still unsettled, largely due to complex morphological variations in *Rorippa*, which were caused by frequent hybridization events. Here, we sequenced four complete plastid genomes of *Rorippa* species by Illumina paired-end sequencing. The four new plastid genomes of *Rorippa* ranged in total size from 154,671 bp for *R. palustris* to 154,894 bp for *R. sylvestris*. There are 130 genes in the four plastomes, embodying 8 rRNA, 37 tRNA, and 85 protein-coding genes. Combining with six published plastid genomes, we carried on comparative and phylogenetic analyses. We found that the ten *Rorippa* plastid genomes were conservative in gene number and order, total size, genomic structure, codon usage, long repeat sequence, and SSR. Fourteen mutational hotspot regions could be selected as candidate DNA barcoding to distinguish *Rorippa* plants. The phylogenetic trees clearly identified that ten *Rorippa* species displayed monophyletic relationships within the tribe Cardamineae based on plastomes and nrDNA ITS sequences. However, there are significant cytonuclear discordances in the interspecific relationships within *Rorippa*, as well as the intergeneric relationships between *Rorippa* and its related genera. We inferred that the cytonuclear discordance is most likely a result of interspecific hybridization within *Rorippa*, as well as intergeneric hybridization with its related genera. These plastid genomes can offer precious information for studies of species authentication, evolutionary history, and the phylogeny of *Rorippa*.

Keywords: *Rorippa*; plastid genome; nrDNA ITS; cytonuclear discordance; DNA barcoding

Citation: Ren, T.; Xun, L.; Jia, Y.; Li, B. Complete Plastomes of Ten *Rorippa* Species (Brassicaceae): Comparative Analysis and Phylogenetic Relationships. *Agronomy* **2024**, *14*, 913. https://doi.org/10.3390/agronomy14050913

Academic Editor: Ainong Shi

Received: 19 March 2024
Revised: 23 April 2024
Accepted: 24 April 2024
Published: 26 April 2024

1. Introduction

The genus *Rorippa* Scopoli, consisting of approximately 75 species, usually has yellow flowers (Figure 1) in the family Brassicaceae [1,2]. It mainly grows in the Northern Hemisphere and occurs on every continent except Antarctica [1]. The taxonomy and phylogenetic relationships of *Rorippa* are still unsettled, largely due to complex morphological variations in *Rorippa*, which were caused by frequent hybridization events [3,4]. The molecular phylogeny of *Rorippa* has been explored by DNA evidence from the *mat*K gene, *rbc*L gene, *trn*L-*trn*F spacer, *trn*T-*trn*L spacer, and *trn*L intron of the plastid genome, as well as the ITS and *Chs* gene of the nuclear genome, while the relationship with its related genera is still unclear [3,5–7]. Perhaps due to a few specific fragments of DNA with limited phylogenetic signals, it therefore heavily inhibited the studies of phylogenetic relationships of *Rorippa*. Currently, with the advent of high-throughput sequencing, it has become comparatively easy to obtain the whole plastid genome, and a growing number of plastomes in Brassicaceae have been published and applied in phylogenomic studies, such as *Cardamine* [8], *Camelina* [9], *Erysimum* [10]. Thus, we tend to utilize the plastid genome to deduce a robust phylogenetic framework of *Rorippa* with respect to Brassicaceae, which not only can clarify the evolutionary relationships of *Rorippa* but can also provide strong evidence for the taxonomic studies of this genus.

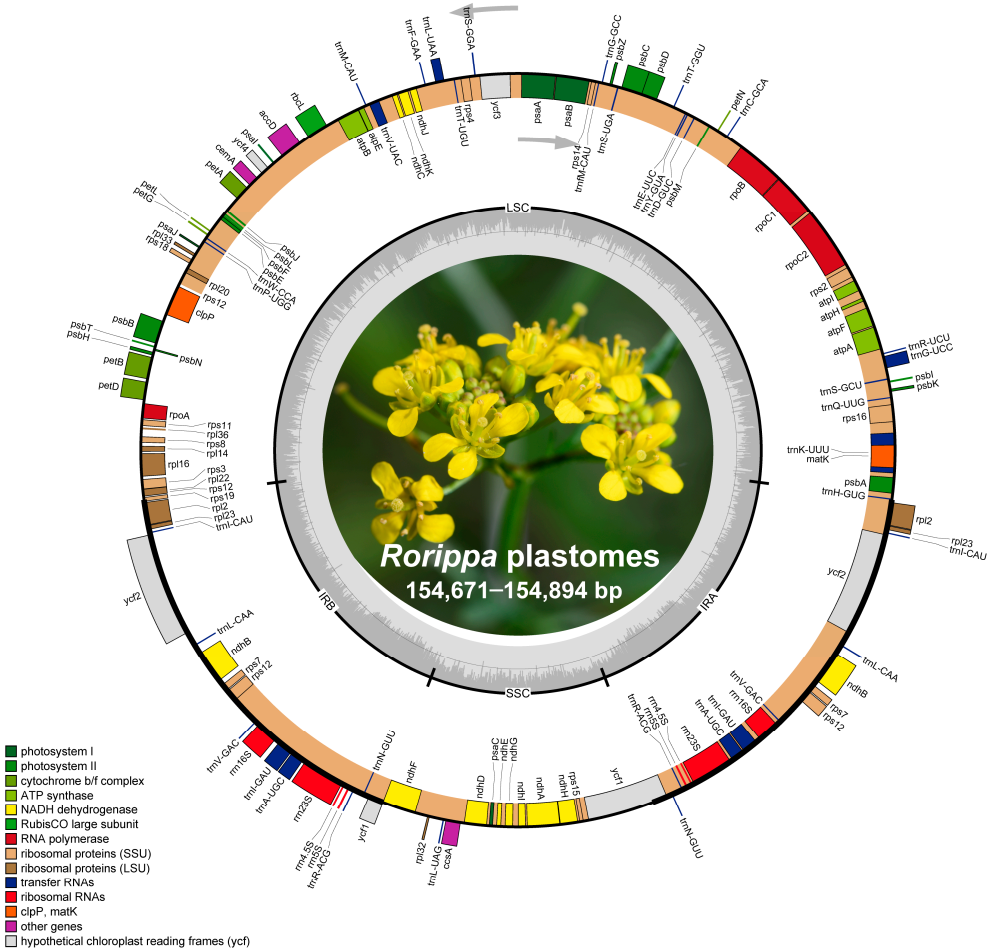

Figure 1. Gene map of four newly sequenced *Rorippa* plastomes. The gray arrows indicate the direction of transcription. The innermost darker gray represents the GC content of the plastome.

The *Rorippa* species has a long history of medicinal use by the Chinese people, owing to the medicinal value of its members. For example, some valuable medicinal components have been found in *R. indica* and several other *Rorippa* plants, including isothiocyanates, glucosinolates, flavonoids, and roripamine [11,12]. Moreover, as a traditional medicine in China, the dried whole-plant material of *R. indica* has the effects of being anti-cough, anti-fever, anti-inflammatory, and having diuretic properties; additionally, it helps blood circulation and rheumatoid arthritis [11]. Recently, owing to the development and utilization of medicinal plants, the identification of wild species is especially important. However, it is difficult to identify the species within the genus *Rorippa* because of hybridization, and morphologically intermediate taxa have been found [13]. As a consequence, exploitation of more discriminating DNA markers for species identification of *Rorippa* is urgently needed in order to guarantee the quality of medicinal materials.

Plastid is a key organelle in plant cells, and it involves in photosynthesis and other biochemical pathways [14]. In most land plants, the plastid genome (plastome) has a relatively conservative circular DNA arrangement with a length of 115–165 kb, embodying four typical regions: small single-copy (SSC) regions of 15–20 kb, two inverted repeats (IRs) of 22–25 kb, and large single-copy (LSC) 82–90 kb [15]. In the plastome, gene content and gene order have been thought to be conserved, normally comprising 110–130 distinct genes [16]. However, rearrangement, large inversions, gene losses, and expansion or contraction of IRs have been documented in their evolution of the plastomes in many angiosperms [17–20]. For instance, the *acc*D, *ccs*A, *clp*P, *inf* A, *ycf* 1, and *ndh* complexes have been lost in some plant lineages [19,21,22]. In addition, the IR regions of some species, such as *Cephalotaxus oliveri* [23], *Taxus chinensis* var. *mairei* [24], *Passiflora* [25], and some species

of Geraniaceae and papilionoid legumes [26,27] showed complete or partial losses. For its slow evolutionary rates, conservative genomic structure, and uniparentally inherited nature of plastomes, the sequences are commonly used as an effective tool for DNA barcoding, evolutionary and phylogenetic studies of plant lineage. For example, the complete plastome sequences as super-barcodes in *Stipa* were much more effective than multi-locus DNA barcodes from plastomes [28]. The phylogenetic relationships and diversification history of Rosaceae were revealed by plastid phylogenomics [29]. A robust molecular phylogeny of Lamiaceae was provided by 79 plastid protein-coding genes and recognized three new tribes [30].

In the present study, we sequenced the plastomes of four *Rorippa* plants (*R. globosa*, *R. indica*, *R. palustris*, and *R. sylvestris*) (Table S1) and conducted an in-depth analysis with previously published six plastomes, which is the first comprehensive analysis of *Rorippa* plastomes. The purposes were (1) to present the newly obtained complete plastid genomes of four *Rorippa* plants; (2) to compare the whole structures of all ten *Rorippa* plastomes; and (3) to improve our understanding of the phylogenetic position of *Rorippa* within Brassicaceae based on plastome sequences.

2. Materials and Methods

2.1. Plant Materials, DNA Extraction, and Sequencing

Four species distributed in China, *R. globosa*, *R. indica*, *R. palustris*, and *R. sylvestris*, were field-collected (The sampling information of four *Rorippa* species in this study was shown in Table S1). Voucher specimens of four *Rorippa* species were stored in the herbarium of the Xi'an Botanical Garden of the Shaanxi Province (XBGH) (Xi'an, China) (Table S1). Fresh and healthy leaves from the *Rorippa* plants were sampled and immediately dried with silica gel. Total genomic DNA was isolated from leaf material according to a modified CTAB method [31] at Novogene (Tianjin, China). The quantities and qualities of genomic DNA were checked on an Agilent BioAnalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). Total genomic DNA was used to construct a sequencing library following the manufacturer's protocol. Paired-end (PE) sequencing libraries with an insert size of 500 bp were sequenced on an Illumina NovaSeq sequencing platform (Illumina, San Diego, CA, USA). Sequencing produced a total of 3.19–3.7 Gb of raw data per species. The complete plastomes of *R. amphibian* (ON411624), *R. cantoniensis* (NC_070424), *R. dubia* (NC_070412), *R. mexicana* (ON892569), *R. sessiliflora* (ON892599), and *R. teres* (ON892567) were recovered to carry on a comparative analysis with these four species.

2.2. Genome Assembly and Annotation

Firstly, clean data were obtained by removing low-quality reads and adaptors using fastP (-n 10 and -q 15) [32]. Then, de novo plastid genome assembly from the clean data was accomplished by GetOrganelle v1.7.2 (-k 21, 45, 65, 85, 105 and -R 15) [33]. Finally, we annotated the complete plastomes of four *Rorippa* species using a GeSeq tool [34], coupled with manually corrected plastid protein-coding genes by Geneious v9.0.2 software (Biomatters Ltd., Auckland, New Zealand) according to its congeneric species. After annotation, the four *Rorippa* plastomes were submitted to the GenBank (PP297065–PP297068). The gene map was constructed by the OrganellarGenomeDRAW tool v1.3.1 [35]. Furthermore, the complete or partial nrDNA (18S-ITS1-5.8S-ITS2-26S) sequences of the four species were also assembled by GetOrganelle v1.7.2 (-k 35, 85, 115 and -R 10) [33] and then uploaded to the GenBank (PP329011–PP329014).

2.3. Codons and Repeat Sequences Analyses

MEGA 6 [36] software was used to analyze the codon usage bias of the plastid protein-coding genes with CDS lengths greater than 300 bp, and the heatmap was produced using TBtools v1.0.0 [37]. The long repetitive sequences comprising palindromic, forward, complementary, and reverse repeats were determined by REPuter [38] (minimum repeat

size = 30 bp and Hamming distance = 3). Perl script MISA [39] was used to detect simple sequence repeats (SSRs) in each species, and the minimum numbers of SSRs were set as ten for mononucleotide repeat, five for dinucleotide repeat, four for trinucleotide repeat, and three for tetra-, penta-, and hexanucleotide repeats.

2.4. Comparative Plastid Genomic Analysis

The IR/SC boundaries of ten *Rorippa* plastomes were compared to elaborate the expansion and contraction of IR regions. The sequence identity of the ten *Rorippa* plastome sequences was analyzed by mVISTA [40]. The plastome regions with an aligned length of over 200 bp were extracted, and the nucleotide diversity (Pi) was then computed with DnaSP v5.1 [41]. Ten plastid genomes rearrangement analysis of *Rorippa* species were conducted by whole genome alignment in Mauve [42].

2.5. Phylogenetic Analysis

We used 54 complete plastomes and corresponding 54 nrDNA ITS sequences from Brassicaceae to explore the phylogenetic position of *Rorippa*. Among them, two *Aethionema* species (*Aethionema cordifolium* and *A. grandiflorum*) were used as the outgroups, while the 76 shared plastid protein-coding genes and 54 nrDNA ITS sequences were used to carry on the phylogenetic analyses. PhyloSuite v1.2.2 [43] was used to extract the 76 shared plastid protein-coding genes. MAFFT v.7 [44] with—auto option was used to obtain the multi-sequence alignments for each gene. TrimAl v. 1.2 [45] with automated1 option was used to trim these alignments. After trimming, the 76 shared plastid protein-coding genes were concatenated using PhyloSuite v1.2.2 [43].

RAxML v8.2.8 [46] was used to conduct the maximum likelihood (ML) analyses with 1000 bootstrap replicates and a GTRGAMMA model. Modeltest v3.7 [47] was used to determine the optimal nucleotide substitution model. The optimal models for plastomes and nrDNA ITS sequences in Bayesian inference (BI) analyses were GTR + I + G and SYM + I + G, respectively. The BI analyses were performed by MrBayes v3.1.2 [48]. Markov chain Monte Carlo (MCMC) runs were initiated with a random tree for 5,000,000 generations sampling every 100 generations, with the first 25% of trees being discarded. Convergence of the MCMC chains was examined when the average standard deviation of split frequencies (ASDSF) was below 0.01.

3. Results

3.1. Plastid Genome Features

The genomic DNA was sequenced by the Illumina NovaSeq sequencing platform yielding the raw data of *R. globosa* (3.19 Gb), *R. indica* (3.7 Gb), *R. palustris* (3.66 Gb), and *R. sylvestris* (3.39 Gb), respectively. The raw data were trimmed by removing low-quality reads and adaptors to obtain the clean data. Additionally, 3,910,034 (*R. globosa*), 3,371,680 (*R. indica*), 4,862,786 (*R. palustris*), and 8,634,650 (*R. sylvestris*) clean reads were used to assemble the four plastomes (Table 1). The coverage depth of the assembled plastome ranged from 408.1× (*R. indica*) to 422.9× (*R. palustris*) (Table 1). All four new *Rorippa* plastomes were similar to other species in Brassicaceae [49]. The four new plastomes of *Rorippa* ranged in total size from 154,671 bp for *R. palustris* to 154,894 bp for *R. sylvestris* (Figure 1; Table 1). The four *Rorippa* plastomes presented a classical quadripartite structure, being composed of two copies of IRs (26,474–26,495 bp) split by an SSC (17,998–18,005 bp) region and an LSC (83,707–83,899 bp) region (Table 1).

Table 1. Basic characteristics of plastomes in ten *Rorippa* species.

	***R. globose* ***	***R. indica* ***	***R. palustris* ***	***R. sylvestris* ***	***R. amphibia***	***R. cantoniensis***	***R. dubia***	***R. mexicana***	***R. sessiliflora***	***R. teres***
Assembly reads	3,910,034	3,371,680	4,862,786	8,634,650	/	/	/	/	/	/
Mean coverage	409.8×	408.1×	422.9×	415.1×	/	/	/	/	/	/
Genome size (bp)	154,675	154,705	154,671	154,894	154,682	155,650	154,740	155,786	155,293	155,095
LSC (bp)	83,711	83,752	83,707	83,899	83,688	84,613	83,741	84,725	84,229	84,067
SSC (bp)	17,998	18,005	17,998	18,005	18,004	18,025	18,015	18,051	17,982	18,012
IR (bp)	26,483	26,474	26,483	26,495	26,495	26,506	26,492	26,505	26,541	26,508
Total GC content (%)	36.4	36.3	36.4	36.4	36.4	36.3	36.4	36.3	36.3	36.3
LSC (%)	34.1	34.1	34.1	34.1	34.1	34.0	34.1	34.0	34.0	34.0
SSC (%)	29.2	29.1	29.2	29.2	29.2	29.2	29.2	29.2	29.3	29.2
IR (%)	42.4	42.3	42.4	42.3	42.3	42.3	42.3	42.3	42.3	42.3
Total gene numbers	130	130	130	130	130	130	130	130	130	130
Protein-coding	85	85	85	85	85	85	84	84	85	85
tRNA	37	37	37	37	37	37	37	37	37	37
rRNA	8	8	8	8	8	8	8	8	8	8
GenBank accession	PP297065	PP297066	PP297067	PP297068	ON411624	NC_070424	NC_070412	ON892569	ON892599	ON892567
References	This study	This study	This study	This study	Genbank	Genbank	Genbank	Genbank	Genbank	Genbank

* The Four Newly Obtained Plastome Sequences.

The plastome sizes, gene number, GC content, and other information of all ten *Rorippa* species are shown in Table 1. The discrepancies between the plastome sizes of the ten *Rorippa* species did not exceed 1115 bp. The overall GC% of the ten plastomes was similar (36.3–36.4%). The GC% was distributed unevenly in different regions of the ten plastomes, displaying as higher in the IR regions (42.32% on average) than in the LSC regions (34.06% on average) and SSC regions (29.2% on average) (Table 1). The plastomes of the ten *Rorippa* species contained 130 genes, embodying 8 rRNA, 37 tRNA, and 84–85 protein-coding genes (Figure 1; Table 1 and Table S2). The *rps*16 gene in *R. dubia* and *R. mexicana* plastome was identified as a pseudogene (Table 1 and Table S2). Additionally, eighteen genes were duplicated, embodying 4 rRNA genes and 14 other genes (Figure 1; Table 1 and Table S2).

3.2. Codon Use Preference Analysis

In total, 21,276–21,302 codons of the 53 CDS genes were encoded in the ten plastomes of *Rorippa* (Table S3). The RSCU value for each species displayed similar codon preference in the 64 codons of the 53 CDSs (Table S3). As a result, 30 of them were used frequently (RSCU > 1); 32 of them were used infrequently (RSCU < 1); and two of them displayed no preferences (RSCU = 1) (Figure 2; Table S3). Among the preferred codons, 29 of them were A/U-ended, except UUG (Figure 2; Table S3). Among the three stop codons, UAA was encoded to be more bountiful than UAG and UGA, thus exhibiting higher preferences. There were no rare codons (RSCU < 0.1) discovered in the CDS genes of the ten *Rorippa* plastomes. Cys was encoded by 243–247 codons (243–247), whereas Leu was encoded by 2242–2247 codons (Table S3).

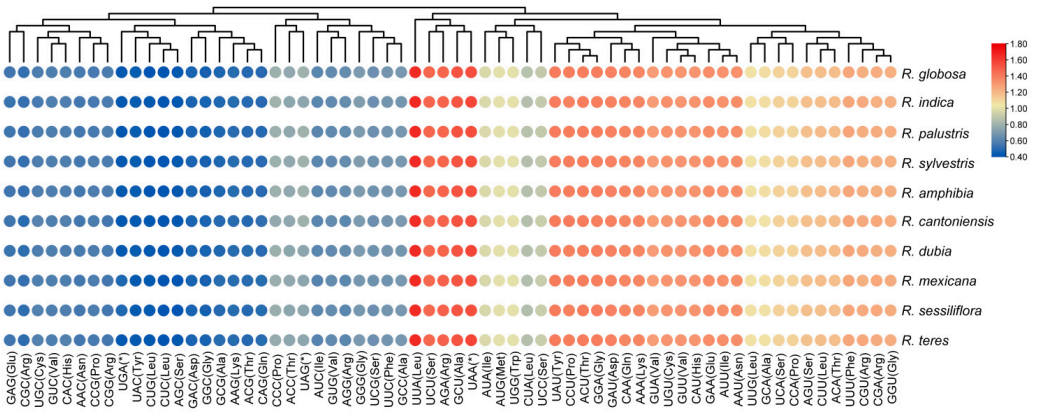

Figure 2. The RSCU values of 53 merged protein-coding sequences for ten *Rorippa* plastomes. Color key: the red values indicate higher RSCU values and the blue values indicate lower RSCU values.

3.3. Long Repeat Sequences and SSR Analyses

Repeat sequences with a length of at least 30 bp in ten *Rorippa* plastomes were detected (Figure 3). There are a total of 303 long repeat sequences composed of 207 forward, 141 palindromic, 12 reverse, and 10 complement repeats (Figure 3; Table S4). Among the ten *Rorippa* plastomes, *R. sylvestris* had the most long repeats (with 45), while *R. teres* had the least long repeats (with 32) (Figure 3A). The long repeat sequences with a length of 30–40 bp were the most common in the ten *Rorippa* plastomes (Figure 3B).

SSRs in the plastomes of ten *Rorippa* species were detected (Figure 4). The number of SSRs in the ten *Rorippa* plastomes was 102 (*R. mexicana*)-111 (*R. sessiliflora*) (Figure 4A; Table S5). The richest SSRs were mononucleotide repeats (748, 71.04%), followed by dinucleotides (188, 17.85%), tetranucleotides (76, 7.22%), and trinucleotides (32, 3.04%). The pentanucleotides and hexanucleotides proved to be very rare across the ten *Rorippa* plastomes (Figure 4A; Table S5). The richest SSRs were A and T nucleotide repeats (such as A/T, AT/AT, AAT/ATT, AAAT/ATTT, AATT/AATT, AAAAT/ATTTT, AATAT/ATATT, and AAATAT/ATATTT motifs), which accounted for 96.11% of the total (Figure 4B).

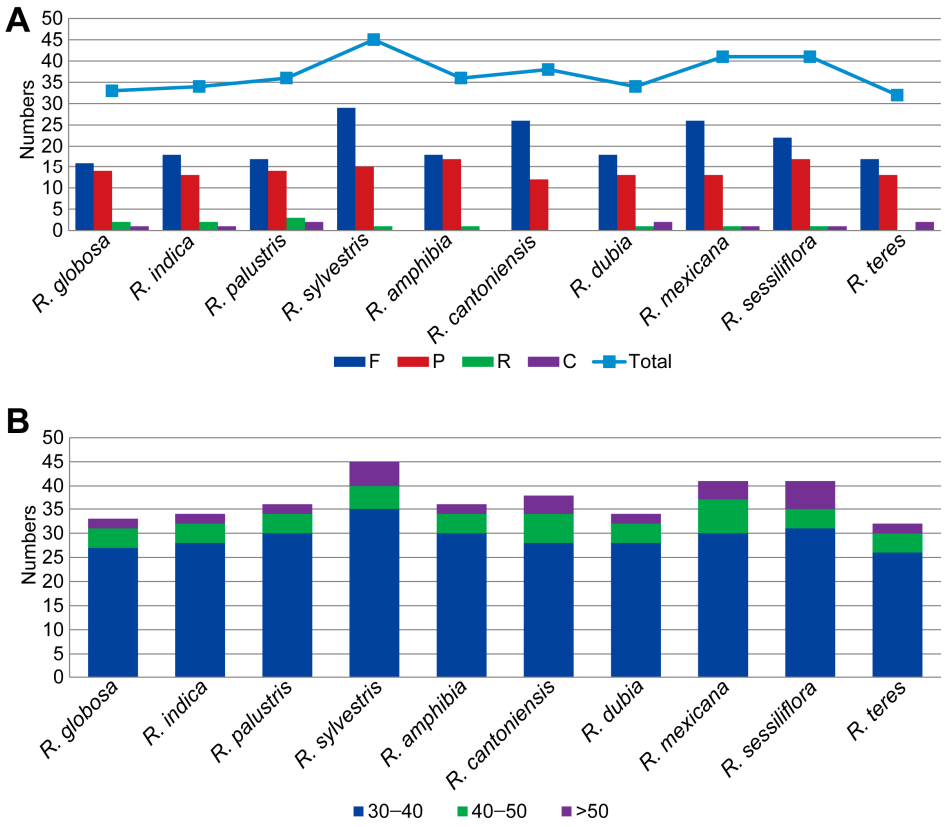

Figure 3. Long repeat sequences in the ten *Rorippa* plastomes. (**A**) Total numbers of four repeat types. (**B**) Number of different repeat lengths. F: forward repeats, P: palindromic repeats, R: reverse repeats, C: complementary repeats.

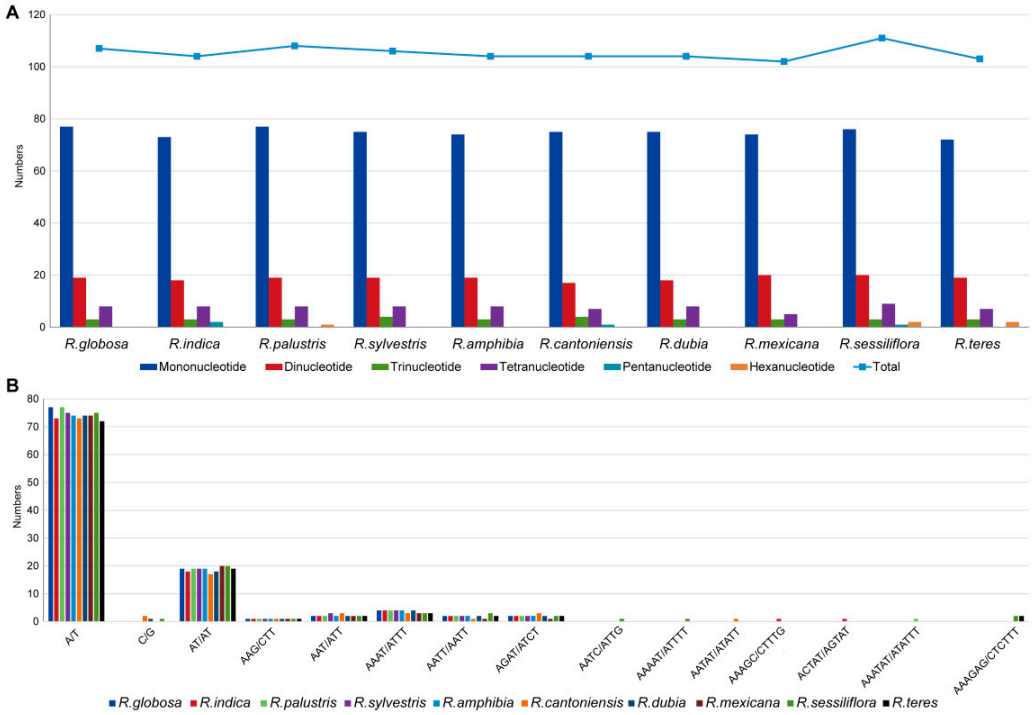

Figure 4. Simple sequence repeat (SSR) in the ten *Rorippa* plastomes. (**A**) Total numbers of SSRs. (**B**) Number of SSR motifs.

3.4. Comparisons of the Plastomes in Rorippa

The ten *Rorippa* plastomes exhibited high levels of structural conservation. The IR regions of the ten *Rorippa* plastomes are the most conserved, ranging from 26,474 bp (*R. indica*) to 26,541 bp (*R. sessiliflora*). The IR boundary regions varied very slightly in ten *Rorippa* species (Figure 5). The LSC/IRb borders expanded 113 bp into *rps*19 gene in all ten *Rorippa* plastomes. The SSC/IRb borders expanded 2 bp or 3 bp into the *ycf*1 genes in the ten *Rorippa* plastomes, whereas the *ndh*F gene overlapped with the SSC/IRb border by 27–37 bp. Spanning the SSC/IRa borders, the *ycf*1 genes were situated in the IRa and SSC regions with 1026–1027 bp and 4352–4382 bp. The *trn*H and *rpl*2 genes were 3 bp and 167 bp away from the IRa/LSC borders in all ten *Rorippa* plastomes. Moreover, no rearrangement occurred in the ten *Rorippa* plastomes (Figure S1).

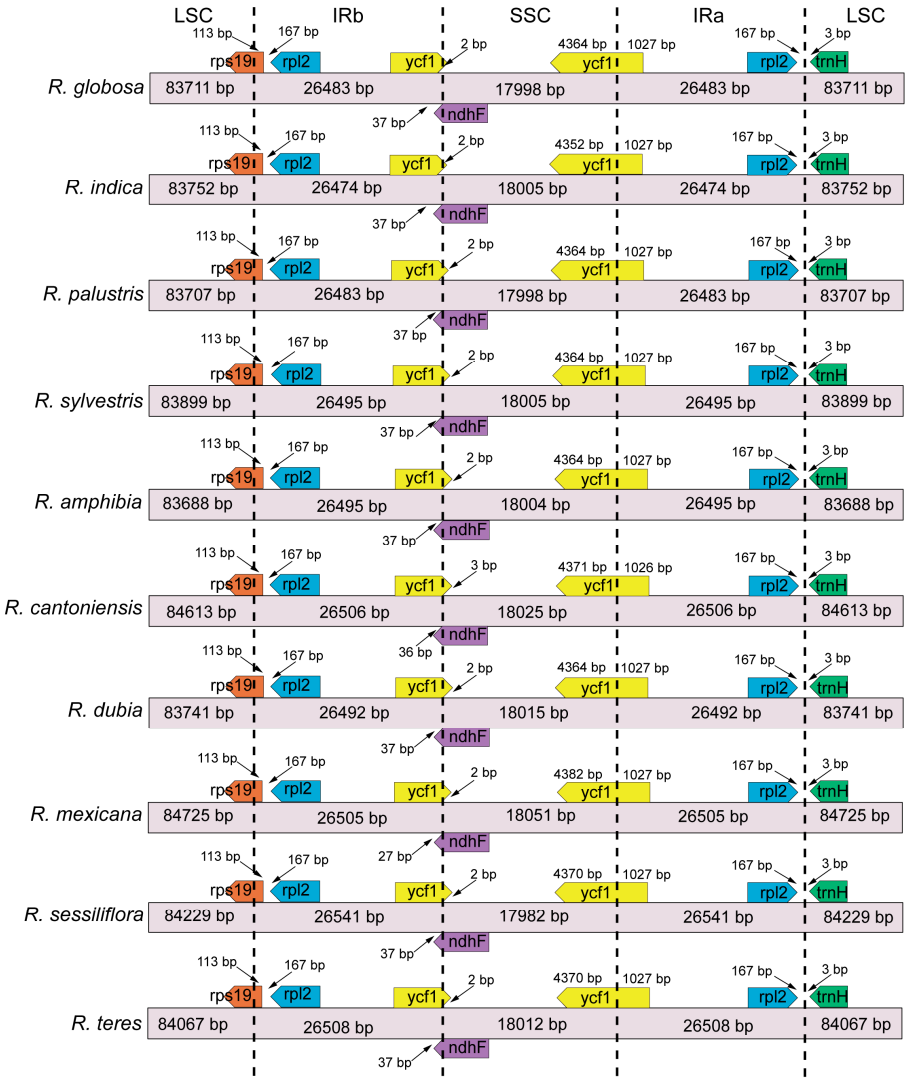

Figure 5. The SC/IR border regions of the ten *Rorippa* plastomes. This figure is not to scale.

3.5. Genomes Sequence Divergence among Rorippa Species

The sequence identity of the ten *Rorippa* plastome sequences was analyzed by using *R. indica* as a reference (Figure S2). The high sequence similarity among the ten plastomes was discovered (Figure S2). Expectantly, non-coding regions and single-copy regions were less conservative than coding regions and IR regions (Figure S2). We also computed the Pi value for 150 regions (Figure 6; Table S6) and obtained the same conclusion as above. In coding regions, the mean Pi value was 0.003758. Due to a higher Pi value (>0.01), two highly variable regions (*mat*K and *ycf*1b) were found, and the Pi value of the *ycf*1 gene

was the highest (0.01146) (Figure 6; Table S6). The mean Pi value of the non-coding regions (including intergenic spacers and introns) (0.011816) was higher than that in the coding regions. We also found twelve highly variable non-coding regions with a higher Pi value (>0.02), namely *ndh*E-*ndh*G, *trn*D-GUC-*trn*Y-GUA, *trn*E-UUC-*trn*T-GGU, *psb*Z-*trn*G-GCC, *trn*K-UUU-*rps*16, *psb*E-*pet*L, *pet*L-*pet*G, *rps*16-*trn*Q-UUG, *psb*K-*psb*I, *rpl*32-*trn*L-UAG, *trn*F-GAA-*ndh*J and *trn*H-GUG-*psb*A, and the Pi value of the *rpl*32-*trn*L-UAG region was highest (0.036942) (Figure 6; Table S6). The mean Pi values in the LSC, IR, and SSC were 0.003852, 0.001177, and 0.005283 in the coding regions, whereas the values in the non-coding regions were 0.0135, 0.001831, and 0.017483, respectively (Table S6).

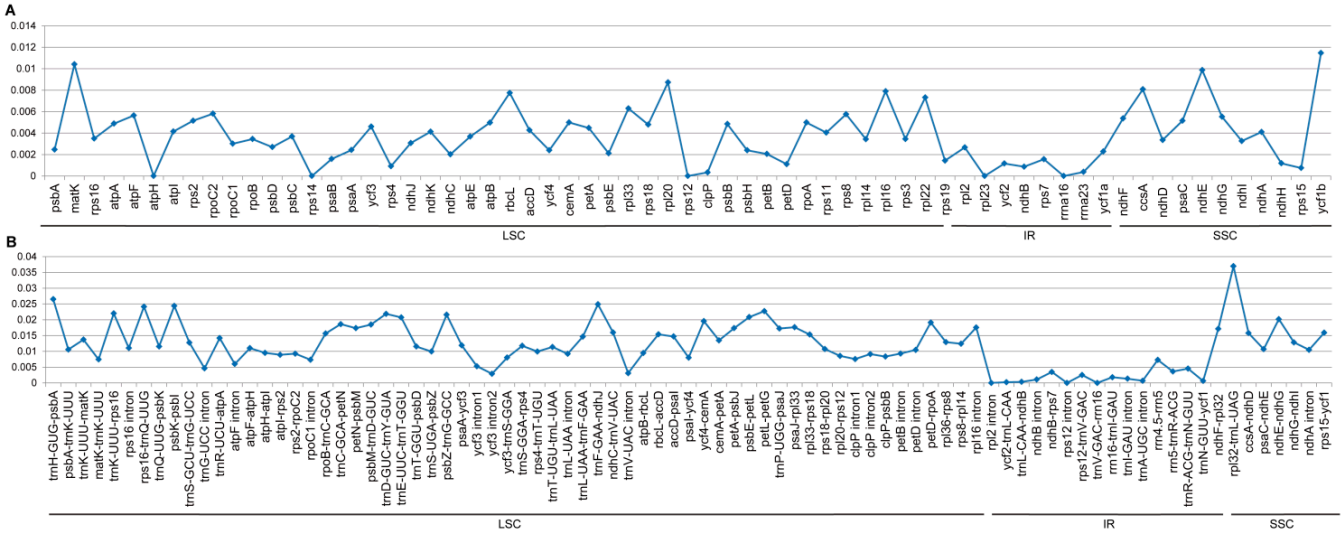

Figure 6. The nucleotide diversity (Pi) of coding (**A**) and non-coding (**B**) regions with an aligned length of over 200 bp within the ten *Rorippa* plastomes.

3.6. Phylogenetic Analysis

We used 54 complete plastomes and corresponding 54 nrDNA ITS sequences from Brassicaceae to explore the phylogenetic position of *Rorippa* (Figure 7; Table S7). *Aethionema cordifolium* and *A. grandiflorum* were used as the outgroups. We found a cytonuclear discordance between the tribes of Brassicaceae in two datasets, which was similar to prior studies [50,51]. As for *Rorippa*, the phylogenetic trees clearly identified that the ten *Rorippa* species displayed monophyletic relationships within the tribe Cardamineae based on the plastomes and nrDNA ITS sequences. However, there is also significant cytonuclear discordance in the intergeneric relationships between *Rorippa* and its related genera, as well as the interspecific relationships within *Rorippa*.

For the phylogenetic analyses of plastid protein-coding sequences, the BI and ML trees exhibited the identical topology (Figure 7A). Almost all of the phylogenetic relationships inferred from 76 shared plastid protein-coding sequences obtained strong support (the range of the support values is 64/0.96–100/1) (Figure 7A). The phylogenetic trees clearly recognized that ten *Rorippa* species were closely related to *Barbarea* (Figure 7A). Two clades were recognized in *Rorippa* with high support (96/1) (Figure 7A). One clade included six *Rorippa* species, among which *R. globosa* and *R. palustris* togethered with *R. amphibia* and *R. sylvestris*, then clustered with *R. indica* and *R. dubia*. Other clades included four *Rorippa* species, among which, *R. teres* and *R. sessiliflora* togethered with *R. mexicana*, then clustered with *R. cantoniensis*.

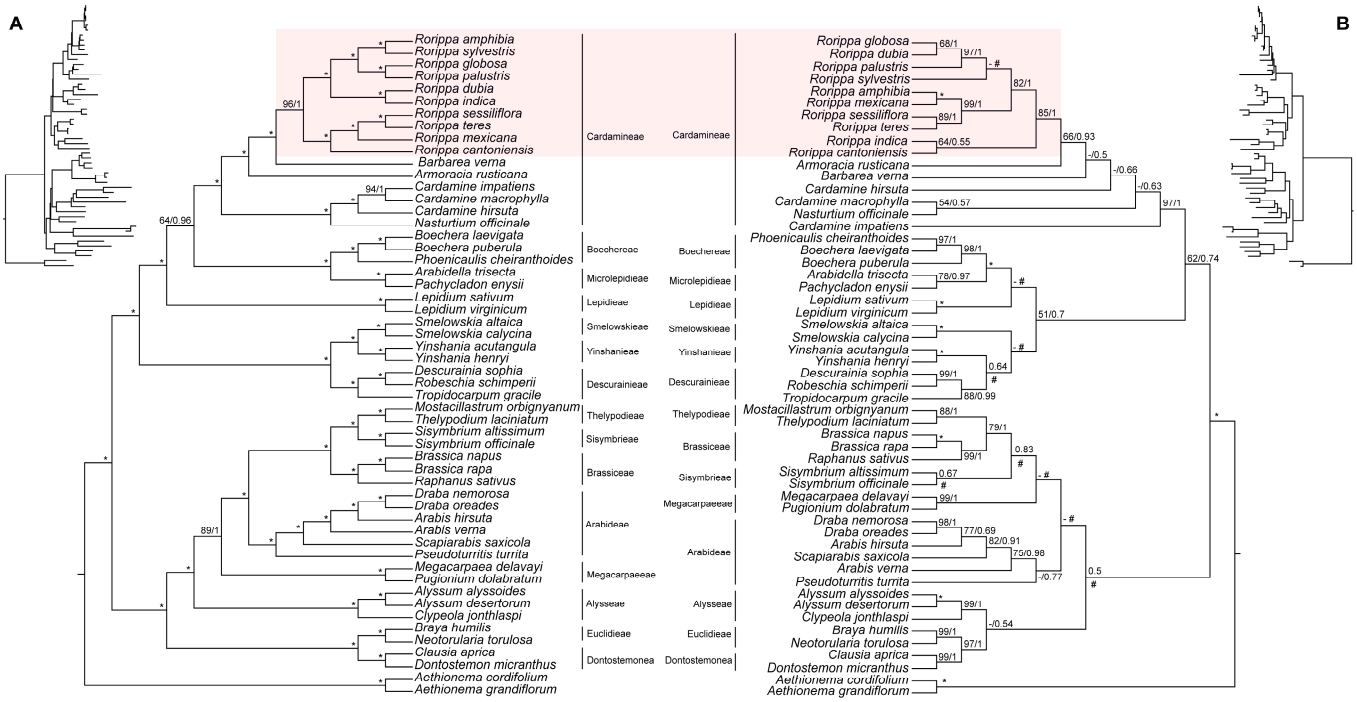

Figure 7. Phylogenetic relationships of *Rorippa* species inferred by 76 shared plastid protein-coding sequences (**A**) and 54 nuclear ribosomal internal transcribed spacer (ITS) sequences (**B**). The numbered-above nodes are Maximum-likelihood bootstrap support values and Bayesian posterior probability values. "*" indicates the highest support values (100%/1), "-" indicates posterior probability values less than 0.5, "#" indicates that the node does not occur in the ML tree.

For the phylogenetic analyses of 54 nrDNA ITS sequences, the tree topologies obtained from BI and ML methods were slightly inconsistent (Figure 7B). The phylogenetic trees identified that ten *Rorippa* species were closely related to *Armoracia* (Figure 7B). Three clades were recognized in *Rorippa* (Figure 7B). One clade included four *Rorippa* species with low support values, among which *R. globosa* and *R. dubia* togethered with *R. palustris* and then clustered with *R. sylvestris*. Other clades included four *Rorippa* species with strong support, among which *R. amphibia* and *R. mexicana* togethered with *R. sessiliflora* and *R. teres*. The remaining two species (*R. indica* and *R. cantoniensis*) formed the third clade.

4. Discussion

4.1. Plastome Evolution of Rorippa

In this study, we obtained four *Rorippa* plastomes and then compared with the six published plastid genomes of *Rorippa*. The ten *Rorippa* plastomes showed conserved gene numbers and orders. All of them contained 130 genes, embodying 8 rRNA, 37 tRNA, and 84–85 protein-coding genes. The *rps*16 gene in *R. dubia* and *R. mexicana* plastome was identified as a pseudogene (it is not uncommon for the *rps*16 gene to be a pseudogene or absent in plant lineage) [22,52]. In Brassicaceae, the *rps*16 gene is in a state of flux with fully functional forms in some species and pseudogenes in others [49,53]. Additionally, the *Rorippa* plastome structure was also highly conserved, and no rearrangement occurred. Like those of most angiosperms, the ten *Rorippa* plastomes have a quadripartite structure, comprising one SSC and LSC, as well as two copies of IR regions [15]. The contraction and expansion of the IR regions often occur in the plastome evolution [54,55]. The LSC/IRb (IRa) borders are completely consistent across the ten *Rorippa* plastomes; meanwhile, the SSC/IRb (IRa) borders undergo slight changes. The discrepancies between the plastome sizes of the ten *Rorippa* species were no greater than 1115 bp. Prior studies showed that the total size of the plastome is often determined by the expansion and contraction of the

IRs [56,57]. Therefore, we speculated that the unconspicuous difference in the total size of the ten *Rorippa* plastomes is owing to the small expansion and contraction of the IR regions.

Codons play a vital role in delivering genetic information, as they are used for translating mRNA into proteins [58]. All the amino acids can be encoded by two or more codons, except for Methionine and Tryptophan. Synonymous codons can encode the same amino acid, while the usage frequency varies in different species [59]. The different usage frequencies of synonymous codons are known as the codon usage bias (CUB). The CUB could be affected by genetic drift, natural selection, and mutation [60,61]. Thus, the study of CUB will promote the understanding of the molecular evolution of the plastome of *Rorippa* species. Ten *Rorippa* plastomes had the same codon usage patterns, embodying 61 amino acid codons and three stop codons. Most of the preferred codons ended in A/U, which is accord with a great many of angiosperms plastomes, such as *Primula* [62], *Phalaenopsis* [63], and *Stephania tetrandra* [64]. The preferred codons usually ending with A/U might be determined by the high AT content in the plastomes [65]. Leu was encoded by the most codons; however, the codon preference order was slightly different from that of *Allium* [66], *Ligusticum* [67], and most Geraniaceae species [26]. In conclusion, the study of codon usage can be deepened by our understanding of the evolutionary history of *Rorippa*.

Long repeat sequences present widely throughout the plastid genome and play essential roles in sequence variation and genome rearrangements [17]. In total, we detected 303 long repeat sequences of four types in the ten plastomes, finding that the number of these repeat types was similar. Among them, the majority of the repeats were 30–40 bp, and palindromic and forward types accounted for the highest proportion, as in previous studies [8,62]. Simple sequence repeats are widely served as molecular markers for population genetic analyses, polymorphism identification, and taxonomic analyses [68]. Here, we identified the SSRs in the ten *Rorippa* species, ranging from 102 to 111, which is comparable to other Brassicaceae in numbers [69,70]. The number of poly (G)/(C) SSRs in the *Rorippa* plastome is significantly less than that of poly (A)/(T), which agrees with the results of other taxa [71]. The cpSSRs identified in the *Rorippa* species are helpful for developing lineage-specific markers for evolution and genetic analyses of this genus.

The mVISTA results showed high sequence similarity across the ten plastomes. The non-coding regions and SC regions were less conservative than coding regions and IR regions, which was confirmed by many Brassicaceae plants [8,72]. The hypervariable regions can serve as potential DNA markers for species identification [73]. Fourteen regions with the highest Pi value have been selected, which might be used as potential cpDNA barcode sequences for *Rorippa* species. Of these, the *mat*K gene with sufficient variant sites has been recognized as a core plant barcode for species discrimination [74]. Some studies have shown that the highly variable *ycf*1 gene can become the DNA barcoding of land plants [75]. The intergenic spacers, such as *psb*K-*psb*I, *psb*E-*pet*L, *trn*E-UUC-*trn*T-GGU, *psb*Z-*trn*G-GCC, *rps*16-*trn*Q-UUG, *rpl*32-*trn*L-UAG, *trn*K-UUU-*rps*16, *trn*F-GAA-*ndh*J, and *trn*H-GUG-*psb*A, have been ascertained in *Quercus* [76], *Rehmannia* [77], Polygonaceae [78], Rhinantheae [79], and *Ligusticum* [67].

4.2. Phylogenetic Relationship of Rorippa

Phylogenetic incongruence between the biparentally inherited nuclear DNA and maternally inherited plastid dataset has been observed in many plant lineages [67,72,80]. Here, we found multiple instances of cytonuclear discordance between our ITS and plastome trees that agree with the results of a recent study [51]. As we all know, there is rampant hybridization in Brassicaceae; therefore, we inferred that cytonuclear discordance between the two datasets is most likely a result of distant hybridization among closely and more distantly related lineages [51].

Likewise, there is also significant cytonuclear discordance in the intergeneric relationships between *Rorippa* and its related genera, as well as the interspecific relationships

within *Rorippa*. Many studies have confirmed that interspecific hybridization has occurred in the genus *Rorippa* [4,13]. Therefore, we inferred that the discordance is most likely a result of interspecific hybridization within *Rorippa* and intergeneric hybridization with its related genera. The phylogenetic trees clearly identified that ten *Rorippa* species displayed monophyletic relationships within the tribe Cardamineae based on plastomes and nrDNA ITS sequences. Furthermore, the tribe Cardamineae also comprises the genera *Barbarea*, *Armoracia*, *Cardamine* and *Nasturtium*, which agrees with previous molecular data [81]. Compared with nrDNA ITS data, our plastome data inferred well-supported relationships of *Rorippa*. The genus *Rorippa* is closely related to *Barbarea* in the plastome tree, which is in accord with the result based on a few DNA marks [6]. Nevertheless, it is hard to compare the interspecific relationships of *Rorippa* species to the previous phylogenetic studies on the limited *Rorippa* species included in this study; therefore, more species should be added to the future phylogenetic studies of *Rorippa*. In addition, due to the frequent hybridization of *Rorippa*, the evolutionary history of the species is complex. Therefore, it is necessary to introduce more nuclear gene data to explore the phylogenetic relationships of *Rorippa*. In short, our study based on plastomes provides a precious resource that should promote the phylogeny, taxonomy, and evolutionary history studies of *Rorippa*.

5. Conclusions

Here, the complete plastid genomes of *R. globosa*, *R. indica*, *R. palustris*, and *R. sylvestris* were assembled and then compared with six published *Rorippa* species. Results of this study showed that the ten *Rorippa* plastomes were conservative in gene number and order, as well as total size, genomic structure, codon usage, long repeat sequence, and SSR. Fourteen mutational hotspot regions (*mat*K, *ycf*1b, *ndh*E-*ndh*G, *trn*D-GUC-*trn*Y-GUA, *trn*E-UUC-*trn*T-GGU, *psb*Z-*trn*G-GCC, *trn*K-UUU-*rps*16, *psb*E-*pet*L, *pet*L-*pet*G, *rps*16-*trn*Q-UUG, *psb*K-*psb*I, *rpl*32-*trn*L-UAG, *trn*F-GAA-*ndh*J, and *trn*H-GUG-*psb*A) could be recognized as candidate DNA barcoding to distinguish *Rorippa* plants. Phylogenetic analyses based on plastid genomes and nrDNA ITS sequences showed that ten *Rorippa* species were monophyletic within the tribe Cardamineae. However, there are significant cytonuclear discordances in the interspecific relationships within *Rorippa*, as well as the intergeneric relationships between *Rorippa* and its related genera. We inferred that these discordances are most likely a result of interspecific hybridization within *Rorippa* and intergeneric hybridization with its related genera.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy14050913/s1, Figure S1: Mauve alignment of ten *Rorippa* plastomes. Within each of the alignments, local collinear blocks are represented by blocks of the same color connected by lines; Figure S2: Sequence identity plot of the ten *Rorippa* plastomes using *R. indica* as a reference; Table S1: Collection locality and voucher information are provided for four newly sequenced plastomes; Table S2: List of genes in the plastome of ten *Rorippa* species; Table S3: Codon usage and relative synonymous codon usage (RSCU) values of protein-coding genes of ten *Rorippa* species; Table S4: Long repeat sequences comparison of ten *Rorippa* species; Table S5: Simple sequence repeats (SSRs) comparison of ten *Rorippa* species; Table S6: Pi values in coding and non-coding regions of ten *Rorippa* plastomes; Table S7: List of species and their accession numbers used for constructing the phylogenetic tree.

Author Contributions: Formal analysis, T.R., L.X. and Y.J.; Funding acquisition, T.R., Y.J. and B.L.; Methodology, T.R.; Project administration, B.L.; Resources, T.R. and L.X.; Software, L.X.; Writing—original draft, T.R.; Writing—review and editing, T.R., Y.J. and B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Project of the Science and Technology Program of Shaanxi Academy of Science (2023K-25, 2023K-48, 2023K-14), National Natural Science Foundation of China (32300314).

Data Availability Statement: Four annotated plastomes and newly sequenced four nrDNA have been submitted into NCBI with accession numbers: PP297065–PP297068 and PP329011–PP329014, respectively.

Acknowledgments: We thank the reviewers who helped improve our manuscript. We also thank the Novogene company for sequencing.

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

References

1. Stuckey, R.L. Taxonomy and distribution of the genus *Rorippa* (Cruciferae) in North America. *SIDA Contrib. Bot.* **1972**, *4*, 279–443.
2. Zhou, T.Y.; Lu, L.L.; Yang, G.; Al-Shehbaz, I.A. *Flora of China*; Science Press: Beijing, China; Missouri Botanical Garden: St. Louis, MO, USA, 2001; Volume 8, pp. 1–193.
3. Bleeker, W.; Weber-Sparenberg, C.; Hurka, H. Chloroplast DNA variation and biogeography in the genus *Rorippa* Scop. (Brassicaceae). *Plant Biol.* **2022**, *4*, 104–111. [CrossRef]
4. Bleeker, W. Interspecific hybridization in *Rorippa* (Brassicaceae): Patterns and processes. *Syst. Biodivers.* **2007**, *5*, 311–319. [CrossRef]
5. Les, D.H. Molecular systematics and taxonomy of lake cress (*Neobeckia aquatica*; Brassicaceae), an imperiled aquatic mustard. *Aquat. Bot.* **1994**, *49*, 149–165. [CrossRef]
6. Koch, M.; Haubold, B.; Mitchell-Olds, T. Molecular systematics of the Brassicaceae: Evidence from coding plastidic *mat*K and nuclear *Chs* sequences. *Am. J. Bot.* **2001**, *88*, 534–544. [CrossRef]
7. Franzke, A.; Pollmann, K.; Bleeker, W.; Kohrt, R.; Hurka, H. Molecular systematics of *Cardamine* and allied genera (Brassicaceae): Its and non-coding chloroplast DNA. *Folia Geobot.* **1998**, *33*, 225–240. [CrossRef]
8. Raman, G.; Park, S. Structural characterization and comparative analyses of the chloroplast genome of Eastern Asian species *Cardamine occulta* (Asian *C. flexuosa* With.) and other *Cardamine* species. *Front. Biosci.* **2022**, *27*, 124. [CrossRef]
9. Brock, J.R.; Mandáková, T.; McKain, M.; Lysak, M.A.; Olsen, K.M. Chloroplast phylogenomics in *Camelina* (Brassicaceae) reveals multiple origins of polyploid species and the maternal lineage of *C. sativa*. *Hortic. Res.* **2022**, *9*, uhab050. [CrossRef] [PubMed]
10. Osuna-Mascaró, C.; Rubio de Casas, R.; Landis, J.B.; Perfectti, F. Genomic resources for *Erysimum* spp. (Brassicaceae): Transcriptome and chloroplast genomes. *Front. Ecol. Evol.* **2021**, *9*, 620601. [CrossRef]
11. Lin, L.Z.; Sun, J.; Chen, P.; Zhang, R.W.; Fan, X.E.; Li, L.W.; Harnly, J.M. Profiling of Glucosinolates and Fla-vonoids in *Rorippa indica* (Linn.) Hiern. (Cruciferae) by UHPLC-PDA-ESI/HRMS^n. *J. Agric. Food Chem.* **2014**, *62*, 6118–6129. [CrossRef]
12. Xu, K.; Chang, Y.; Zhang, Y.; Liu, K.; Zhang, J.; Wang, W.; Li, Z.; Ma, S.; Xin, Y.; Li, C.; et al. *Rorippa indica* regeneration via somatic embryogenesis involving frog egg-like bodies efficiently induced by the synergy of salt and drought stresses. *Sci. Rep.* **2016**, *6*, 19811. [CrossRef] [PubMed]
13. Bleeker, W.; Hurka, H. Introgressive hybridization in *Rorippa* (Brassicaceae): Gene flow and its consequences in natural and anthropogenic habitats. *Mol. Ecol.* **2001**, *10*, 2013–2022. [CrossRef]
14. Raven, J.A.; Allen, J.F. Genomics and chloroplast evolution: What did cyanobacteria do for plants? *Genome Biol.* **2003**, *4*, 1–5. [CrossRef] [PubMed]
15. Ravi, V.; Khurana, J.P.; Tyagi, A.K.; Khurana, P. An update on chloroplast genomes. *Plant Syst. Evol.* **2008**, *271*, 101–122. [CrossRef]
16. Jansen, R.K.; Raubeson, L.A.; Boore, J.L.; Chumley, T.W.; Haberle, R.C.; Wyman, S.K. Methods for obtaining and analyzing whole chloroplast genome sequences. *Methods Enzymol.* **2005**, *395*, 348–384. [PubMed]
17. Weng, M.L.; Blazier, J.C.; Govindu, M.; Jansen, R.K. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats and nucleotide substitution rates. *Mol. Biol. Evol.* **2014**, *31*, 645–659. [CrossRef]
18. Kim, Y.; Cullis, C. A novel inversion in the chloroplast genome of marama (*Tylosema esculentum*). *J. Exp. Bot.* **2017**, *68*, 2065–2072. [CrossRef] [PubMed]
19. Yao, G.; Jin, J.J.; Li, H.T.; Yang, J.B.; Mandala, V.S.; Croley, M.; Mostow, R.; Douglas, N.A.; Chase, M.W.; Christenhusz, M.J.M.; et al. Plastid phylogenomic insights into the evolution of Caryophyllales. *Mol. Phylogenet. Evol.* **2019**, *134*, 74–86. [CrossRef]
20. Ren, T.; Xie, D.; Peng, C.; Gui, L.; Price, M.; Zhou, S.; He, X. Molecular evolution and phylogenetic relationships of *Ligusticum* (Apiaceae) inferred from the whole plastome sequences. *BMC Ecol. Evol.* **2022**, *22*, 55. [CrossRef]
21. Liu, L.X.; Li, R.; Worth, J.R.; Li, X.; Li, P.; Cameron, K.M.; Fu, C.X. The complete chloroplast genome of Chinese bayberry (*Morella rubra*, Myricaceae): Implications for understanding the evolution of Fagales. *Front. Plant Sci.* **2017**, *8*, 264170. [CrossRef]
22. Mohanta, T.K.; Mishra, A.K.; Khan, A.; Hashem, A.; Abd_Allah, E.F.; Al-Harrasi, A. Gene loss and evolution of the plastome. *Genes* **2020**, *11*, 1133. [CrossRef] [PubMed]
23. Yi, X.; Gao, L.; Wang, B.; Su, Y.J.; Wang, T. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): Evolutionary comparison of *Cephalotaxus* chloroplast DNAs and insights into the loss of inverted repeat copies in Gymnosperms. *Genome Biol. Evol.* **2013**, *5*, 688–698. [CrossRef]
24. Zhang, Y.Z.; Ma, J.; Yang, B.X.; Li, R.Y.; Zhu, W.; Sun, L.L.; Tian, J.K.; Zhang, L. The complete chloroplast genome sequence of *Taxus chinensis* var. *mairei* (Taxaceae): Loss of an inverted repeat region and comparative analysis with related species. *Gene* **2014**, *540*, 201–209. [PubMed]

25. Cauz-Santos, L.A.; Da Costa, Z.P.; Callot, C.; Cauet, S.; Zucchi, M.I.; Bergès, H.; Van den Berg, C.; Vieira, M.L.C. A repertory of rearrangements and the loss of an inverted repeat region in *Passiflora* chloroplast genomes. *Genome Biol. Evol.* **2020**, *12*, 1841–1857. [CrossRef] [PubMed]
26. Guisinger, M.M.; Kuehl, J.V.; Boore, J.L.; Jansen, R.K. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: Rearrangements, repeats, and codon usage. *Mol. Biol. Evol.* **2011**, *28*, 583–600. [CrossRef] [PubMed]
27. Choi, I.S.; Jansen, R.; Ruhlman, T. Lost and Found: Return of the inverted repeat in the legume clade defined by its absence. *Genome Biol. Evol.* **2019**, *11*, 1321–1333. [CrossRef] [PubMed]
28. Krawczyk, K.; Nobis, M.; Myszczyński, K.; Klichowska, E.; Sawicki, J. Plastid super-barcodes as a tool for species discrimination in feather grasses (Poaceae: *Stipa*). *Sci. Rep.* **2018**, *8*, 1924. [CrossRef]
29. Zhang, S.D.; Jin, J.J.; Chen, S.Y.; Chase, M.W.; Soltis, D.E.; Li, H.T.; Yang, J.B.; Li, D.Z.; Yi, T.S. Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytol.* **2017**, *214*, 1355–1367. [CrossRef] [PubMed]
30. Zhao, F.; Chen, Y.P.; Salmaki, Y.; Drew, B.T.; Wilson, T.C.; Scheen, A.C.; Celep, F.; Bräuchler, C.; Bendiksby, M.; Wang, Q.; et al. An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Biol.* **2021**, *19*, 2. [CrossRef]
31. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem. Bull.* **1987**, *19*, 11–15.
32. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef] [PubMed]
33. Jin, J.J.; Yu, W.B.; Yang, J.B.; Song, Y.; Depamphilis, 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]
34. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq-Versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [CrossRef]
35. Greiner, S.; Lehwark, P.; Bock, R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **2019**, *47*, W59–W64. [CrossRef] [PubMed]
36. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef]
37. Chen, C.J.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.H.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of bigbiological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]
38. 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]
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. 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]
41. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
42. Darling, A.C.E.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394–1403. [CrossRef]
43. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [CrossRef] [PubMed]
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] [PubMed]
45. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef] [PubMed]
46. Stamatakis, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analysis with thousands of taxa and mixed models. *Bioinformatics* **2006**, *22*, 2688–2690. [CrossRef] [PubMed]
47. Posada, D.; Crandall, K.A. Modeltest: Testing the model of DNA substitution. *Bioinformatics* **1998**, *14*, 817–818. [CrossRef] [PubMed]
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. Guo, X.; Liu, J.; Hao, G.; Zhang, L.; Mao, K.; Wang, X.; Zhang, D.; Ma, T.; Hu, Q.; Al-Shehbaz, I.A.; et al. Plastome phylogeny and early diversification of Brassicaceae. *BMC Genom.* **2017**, *18*, 176. [CrossRef]
50. Mabry, M.E.; Brose, J.M.; Blischak, P.D.; Sutherland, B.; Dismukes, W.T.; Bottoms, C.A.; Edger, P.P.; Washburn, J.D.; An, H.; Hall, J.C.; et al. Phylogeny and multiple independent whole-genome duplication events in the Brassicales. *Am. J. Bot.* **2020**, *107*, 1148–1164. [CrossRef]
51. Hendriks, K.P.; Kiefer, C.; Al-Shehbaz, I.A.; Bailey, C.D.; van Huysduynen, A.H.; Nikolov, L.A.; Nauheimer, L.; Zuntini, A.R.; German, D.A.; Franzke, A.; et al. Global Brassicaceae phylogeny based on filtering of 1,000-gene dataset. *Curr. Biol.* **2023**, *33*, 4052–4068. [CrossRef]

52. Schwarz, E.N.; Ruhlman, T.A.; Sabir, J.S.M.; Hajirah, N.H.; Alharbi, N.S.; Al-Malki, A.L.; Bailey, C.D.; Jansen, R.K. Plastid genome sequences of legumes reveal parallel inversions and multiple losses of *rps*16 in papilionoids. *J. Syst. Evol.* **2015**, *53*, 458–468. [CrossRef]
53. Roy, S.; Ueda, M.; Kadowaki, K.; Tsutsumi, N. Different status of the gene for ribosomal protein S16 in the chloroplast genome during evolution of the genus *Arabidopsis* and closely related species. *Genes Genet. Syst.* **2010**, *85*, 319–326. [CrossRef] [PubMed]
54. Kim, K.J.; Lee, H.L. Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Res.* **2004**, *11*, 247–261. [CrossRef] [PubMed]
55. Kaila, T.; Chaduvla, P.K.; Saxena, S.; Bahadur, K.; Gahukar, S.J.; Chaudhury, A.; Sharma, T.R.; Singh, N.K.; Gaikwad, K. Chloroplast genome sequence of Pigeonpea (*Cajanus cajan* (L.) Millspaugh) and *Cajanus scarabaeoides* (L.) Thouars: Genome organization and comparison with other Legumes. *Front. Plant Sci.* **2016**, *7*, 1847. [CrossRef] [PubMed]
56. Green, B.R. Chloroplast genomes of photosynthetic eukaryotes. *Plant J.* **2011**, *66*, 34–44. [CrossRef] [PubMed]
57. Weng, M.L.; Ruhlman, T.A.; Jansen, R.K. Expansion of inverted repeat does not decrease substitution rates in *Pelargonium* plastid genomes. *New Phytol.* **2017**, *214*, 842–851. [CrossRef] [PubMed]
58. Qiu, S.; Zeng, K.; Slotte, T.; Wright, S.; Charlesworth, D. Reduced efficacy of natural selection on codon usage bias in selfing *Arabidopsis* and *Capsella* species. *Genome Biol. Evol.* **2011**, *3*, 868–880. [CrossRef]
59. Grantham, R.; Gautier, C.; Gouy, M. Codon frequencies in 119 individual genes confirm corsistent choices of degenerate bases according to genome type. *Nucleic Acids Res.* **1980**, *8*, 1893–1912. [CrossRef]
60. Akashi, H. Codon bias evolution in *Drosophila*. Population genetics of mutation-selection drift. *Gene* **1997**, *205*, 269–278. [CrossRef]
61. Chen, S.L.; Lee, W.; Hottes, A.K.; Shapiro, L.; McAdams, H.H. Codon usage between genomes is constrained by genome-wide mutational processes. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3480–3485. [CrossRef]
62. Ren, T.; Yang, Y.C.; Zhou, T.; Liu, Z.L. Comparative plastid genomes of *Primula* species: Sequence divergence and phylogenetic relationships. *Int. J. Mol. Sci.* **2018**, *19*, 1050. [CrossRef] [PubMed]
63. Tao, L.; Duan, H.; Tao, K.; Luo, Y.; Li, Q.; Li, L. Complete chloroplast genome structural characterization of two *Phalaenopsis* (Orchidaceae) species and comparative analysis with their alliance. *BMC Genom.* **2023**, *24*, 359. [CrossRef] [PubMed]
64. Dong, S.; Ying, Z.; Yu, S.; Wang, Q.; Liao, G.; Ge, Y.; Cheng, R. Complete chloroplast genome of *Stephania tetrandra* (Menispermaceae) from Zhejiang Province: Insights into molecular structures, comparative genome analysis, mutational hotspots and phylogenetic relationships. *BMC Genom.* **2021**, *22*, 880. [CrossRef]
65. Tian, S.; Lu, P.; Zhang, Z.; Wu, J.Q.; Zhang, H.; Shen, H. Chloroplast genome sequence of Chongming lima bean (*Phaseolus lunatus* L.) and comparative analyses with other legume chloroplast genomes. *BMC Genom.* **2021**, *22*, 194. [CrossRef] [PubMed]
66. Cheng, R.Y.; Xie, D.F.; Zhang, X.Y.; Fu, X.; He, X.J.; Zhou, S.D. Comparative plastome analysis of three Amaryllidaceae subfamilies: Insights into variation of genome characteristics, phylogeny, and adaptive evolution. *BioMed Res. Int.* **2022**, *2022*, 3909596. [CrossRef] [PubMed]
67. Ren, T.; Li, Z.X.; Xie, D.F.; Gui, L.J.; Peng, C.; Wen, J.; He, X.J. Plastomes of eight *Ligusticum* species: Characterization, genome evolution, and phylogenetic relationships. *BMC Plant Biol.* **2020**, *20*, 519. [CrossRef] [PubMed]
68. Xue, J.; Wang, S.; Zhou, S.L. Polymorphic chloroplast microsatellite loci in *Nelumbo* (Nelumbonaceae). *Am. J. Bot.* **2012**, *99*, e240–e244. [CrossRef]
69. Ren, T.; Peng, F.F.; Wei, Y.F.; Luo, S.X.; Liu, Z.L. Characterization of the complete plastid genome of *Draba oreades* (Brassicaceae). *Mitochondrial DNA B Resour.* **2019**, *4*, 439–440. [CrossRef]
70. Huang, S.; Kang, Z.; Chen, Z.; Deng, Y. Comparative analysis of the chloroplast genome of *Cardamine hupingshanensis* and phylogenetic study of *Cardamine*. *Genes* **2022**, *13*, 2116. [CrossRef] [PubMed]
71. Zhao, K.; Li, L.; Quan, H.; Yang, J.; Zhang, Z.; Liao, Z.; Lan, X. Comparative analyses of chloroplast genomes from 14 *Zanthoxylum* species: Identification of variable DNA markers and phylogenetic relationships within the genus. *Front. Plant Sci.* **2021**, *11*, 605793. [CrossRef]
72. Hu, H.; Hu, Q.; Al-Shehbaz, I.A.; Luo, X.; Zeng, T.; Guo, X.; Liu, J. Species delimitation and interspecific relationships of the genus *Orychophragmus* (Brassicaceae) inferred from whole chloroplast genomes. *Front. Plant Sci.* **2016**, *7*, 1826. [CrossRef] [PubMed]
73. Nock, C.J.; Waters, D.L.; Edwards, M.A.; Bowen, S.G.; Rice, N.; Cordeiro, G.M.; Henry, R.J. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol. J.* **2011**, *9*, 328–333. [CrossRef] [PubMed]
74. CBOL Plant Working Group. A DNA barcode for land plants. *Proc. Nat. Acad. Sci. USA* **2009**, *106*, 12794–12797. [CrossRef]
75. Dong, W.P.; Xu, C.; Li, C.H.; Sun, J.H.; Zuo, Y.J.; Shi, S.; Cheng, T.; Guo, J.J.; Zhou, S.L. *ycf*1, the most promising plastid DNA barcode of land plants. *Sci. Rep.* **2015**, *5*, 8348. [CrossRef] [PubMed]
76. Yang, Y.C.; Zhou, T.; Duan, D.; Yang, J.; Feng, L.; Zhao, G.F. Comparative analysis of the complete chloroplast genomes of five *Quercus* species. *Front. Plant Sci.* **2016**, *7*, 959. [CrossRef] [PubMed]
77. Zeng, S.; Zhou, T.; Han, K.; Yang, Y.; Zhao, J.; Liu, Z.L. The complete chloroplast genome sequences of six *Rehmannia* species. *Genes* **2017**, *8*, 103. [CrossRef] [PubMed]
78. Wang, X.; Zhou, T.; Bai, G.; Zhao, Y. Complete chloroplast genome sequence of *Fagopyrum dibotrys*: Genome features, comparative analysis and phylogenetic relationships. *Sci. Rep.* **2018**, *8*, 12379. [CrossRef] [PubMed]
79. Zhou, T.; Wang, J.; Li, W.; Zhang, X.; Xu, Y.; Xu, F.; Zhu, H.; Wang, X. The complete chloroplast genome of *Euphrasia regelii*, pseudogenization of *ndh* genes and the phylogenetic relationships within Orobanchaceae. *Front. Genet.* **2019**, *10*, 444. [CrossRef]

80. Ji, Y.; Liu, C.; Yang, Z.; Yang, L.; He, Z.; Wang, H.; Yang, J.; Yi, T. Testing and using complete plastomes and ribosomal DNA sequences as the next generation DNA barcodes in *Panax* (Araliaceae). *Mol. Ecol. Resour.* **2019**, *19*, 1333–1345. [CrossRef]
81. Carlsen, T.; Bleeker, W.; Hurka, H.; Elven, R.; Brochmann, C. Biogeography and phylogeny of *Cardamine* (Brassicaceae). *Ann. Mo. Bot. Gard.* **2009**, *96*, 215–236. [CrossRef]