



# Complete Chloroplast Genomes Provide Insights Into Evolution and Phylogeny of *Campylotropis* (Fabaceae)

Yu Feng<sup>1</sup>, Xin-Fen Gao<sup>1,2</sup>, Jun-Yi Zhang<sup>1</sup>, Li-Sha Jiang<sup>1,2</sup>, Xiong Li<sup>1,2</sup>, Heng-Ning Deng<sup>1</sup>, Min Liao<sup>1,2</sup> and Bo Xu<sup>1,2,3\*</sup>

<sup>1</sup>CAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization and Ecological Restoration and Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China, <sup>2</sup>College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, <sup>3</sup>Mangkang Ecological Station, Tibet Ecological Safety Monitor Network, Changdu, China

## OPEN ACCESS

### Edited by:

Ruslan Kalendar,  
University of Helsinki, Finland

### Reviewed by:

Revathi Sekar,  
Helmholtz Association of German  
Research Centres (HZ), Germany  
M. Gonzalo Claros,  
University of Malaga, Spain  
Dario Ojeda Alayon,  
Norwegian Institute of Bioeconomy  
Research, Norway

### \*Correspondence:

Bo Xu  
xubo@cib.ac.cn

### Specialty section:

This article was submitted to  
Plant Systematics and Evolution,  
a section of the journal  
Frontiers in Plant Science

Received: 14 March 2022

Accepted: 14 April 2022

Published: 18 May 2022

### Citation:

Feng Y, Gao X-F, Zhang J-Y,  
Jiang L-S, Li X, Deng H-N,  
Liao M and Xu B (2022) Complete  
Chloroplast Genomes Provide  
Insights Into Evolution and Phylogeny  
of *Campylotropis* (Fabaceae).  
*Front. Plant Sci.* 13:895543.  
doi: 10.3389/fpls.2022.895543

The genus *Campylotropis* Bunge (Desmodieae, Papilionoideae) comprises about 37 species distributed in temperate and tropical Asia. Despite the great potential in soil conservation, horticulture, and medicine usage, little is known about the evolutionary history and phylogenetic relationships of *Campylotropis* due to insufficient genetic resources. Here, we sequenced and assembled 21 complete chloroplast genomes of *Campylotropis* species. In combination with the previously published chloroplast genomes of *C. macrocarpa* and closely related species, we conducted comparative genomics and phylogenomic analysis on these data. Comparative analysis of the genome size, structure, expansion and contraction of inverted repeat (IR) boundaries, number of genes, GC content, and pattern of simple sequence repeats (SSRs) revealed high similarities among the *Campylotropis* chloroplast genomes. The activities of long sequence repeats contributed to the variation in genome size and gene content in *Campylotropis* chloroplast genomes. The *Campylotropis* chloroplast genomes showed moderate sequence variation, and 13 highly variable regions were identified for species identification and further phylogenetic studies. We also reported one more case of *matK* pseudogene in the legume family. The phylogenetic analysis confirmed the monophyly of *Campylotropis* and the sister relationship between *Lespedeza* and *Kummerowia*, the latter two genera were then sister to *Campylotropis*. The intrageneric relationships of *Campylotropis* based on genomic scale data were firstly reported in this study. The two positively selected genes (*atpF* and *rps19*) and eight fast-evolving genes identified in this study may help us to understand the adaptation of *Campylotropis* species. Overall, this study enhances our understanding of the chloroplast genome evolution and phylogenetic relationships of *Campylotropis*.

**Keywords:** *Campylotropis*, legume, adaptive evolution, phylogenomics, comparative genomics, chloroplast genome

## INTRODUCTION

The genus *Campylotropis* Bunge belongs to the tribe Desmodieae (Benth.) Hutchinson in the legume subfamily Papilionoideae. It comprises *c.* 37 species of deciduous shrubs and subshrubs that distributed in Asia from the Himalaya region through Southeast Asia to China and Korea (Barham, 1997; Iokawa and Ohashi, 2008; Huang et al., 2010). Southwest China is the diversity center of *Campylotropis* as it harbors *c.* 80% of the species, and *c.* 20 species are endemic to this region (Iokawa and Ohashi, 2008). Most species in this genus have important value in soil conservation due to their tolerance of arid soils (Huang et al., 2010). Some *Campylotropis* species are also valuable for horticulture and medicine usage. For example, *C. polyantha* is widely used in gardening due to its numerous racemes of showy flowers and long-lasting fluorescence (Barham, 1997). The dried roots of *C. hirtella* can be used as traditional Chinese medicine for the treatment of benign prostate hyperplasia (Wen et al., 2007), and *C. trigonoclada* contains daucosterol linoleate which can be used for the treatment of breast cancer (Han et al., 2018).

As suggested by previous molecular phylogenetic studies, *Campylotropis* is sister to the other two genera of subtribe Lespedezinae (i.e., *Lespedeza* and *Kummerowia*) in tribe Desmodieae (Xu et al., 2012; Jabbour et al., 2018; Jin et al., 2019). Much effort has been made to clarify species relationships within *Campylotropis*, mostly based on morphological characters such as leaf and calyx morphology (e.g., Iokawa and Ohashi, 2008; Huang et al., 2010). However, most of the morphological characteristics (e.g., persistence of bracts, the color of flowers, and shape of keel petals) are polymorphic and vary continuously among species, causing controversial species delimitation in this genus (Iokawa and Ohashi, 2008). Besides, little is known about its intrageneric and interspecific relationships due to the lack of comprehensive molecular phylogenetic studies.

Chloroplasts, derived from photosynthetic bacteria, play critical roles in the survival, adaptation, and evolution of plants (Wicke et al., 2011; Zhao et al., 2019; Dopp et al., 2021). Although the chloroplast (cp) genomes are much smaller than most nuclear genomes, they encode essential proteins related to photosynthesis, fixation of carbon and nitrogen, and biosynthesis of starch, pigments, fatty acids, and amino acids (Howe et al., 2003; Wicke et al., 2011; Daniell et al., 2016). Chloroplast genomes have relatively stable structure and gene content compared to nuclear genomes. The typical structure of angiosperm cp genome is a circular double-stranded DNA molecule, exhibiting a conserved quadripartite structure [i.e., two inverted repeats (IRs) separated by a large single-copy region (LSC) and a small single-copy region (SSC)] and containing 110–130 genes (Sugiura, 1992; Daniell et al., 2016). The characteristics of cp genomes including lack of recombination, low nucleotide substitution rates, and usually uniparental inheritance make them the primary source to explore phylogenetic evolution of plant species (Shaw et al., 2005). Besides, structural variants such as expansion and contraction of IRs, gains or losses of genes and introns, and dynamics of repeat sequences (e.g., simple sequence repeat, SSR) provide resources for

evaluating genomic evolutionary history (e.g. Sabir et al., 2014; Keller et al., 2017). The development of sequencing technology and analysis tools makes the acquisition of cp genomes much easier than before, thus promptly extending gene-based phylogenetics to phylogenomics (Lu et al., 2017). In fact, recent phylogenomic studies have been successful in reconstructing phylogenies at various taxonomic scales (e.g., genera and families) across angiosperms using the cp genome datasets (e.g., Cai et al., 2015; Ruhsam et al., 2015; Luo et al., 2016; Zhang et al., 2017, 2021).

Here, we present 21 complete cp genomes of *Campylotropis* species assembled from Illumina short reads. In combination with the previously published cp genomes of *C. macrocarpa* (Jin et al., 2019) and closely related species, we conducted comparative genomics and phylogenomic analyses on these data with the following aims: (1) to reveal the global structural patterns of *Campylotropis* cp genomes; (2) to investigate variations of SSRs and repeat sequences among *Campylotropis* cp genomes; (3) to screen highly variable regions suitable for species identification and phylogenetic studies; (4) to reconstruct a robust phylogenetic relationship within *Campylotropis* and among genera in the tribe Desmodieae; and (5) to investigate adaptive evolution patterns of cp genes in *Campylotropis*. These results will provide insights into the evolutionary history of *Campylotropis* and tribe Desmodieae as well as abundant information for future phylogenetic and population genetic studies.

## MATERIALS AND METHODS

### Taxon Sampling, DNA Extraction, and Sequencing

In this study, leaf materials of 21 accessions representing 17 *Campylotropis* species (including four subspecies, one variety, and one forma) were collected from the field and preserved in silica gel (Table 1). Voucher specimens were deposited in the Herbarium of the Chengdu Institute of Biology (CDBI; Supplementary Table S1). The extraction of total genomic DNA, library preparation, and Illumina sequencing for each accession were described in our previous study (Liao et al., 2021).

### Chloroplast Genome Assembly, Annotation, and Comparison

For each accession, ~25 Gb of raw data were generated with pair-end 150bp read length. Trimmomatic v0.39 (Bolger et al., 2014) was used to remove low-quality and adapter-containing reads. The clean data were then assembled using GetOrganelle v1.7.5 (Jin et al., 2020). Plastid Genome Annotator (Qu et al., 2019) was used to annotate the cp genomes based on one published accession of *Campylotropis* (*C. macrocarpa*; NC\_044100; Jin et al., 2019) and 15 accessions of closely related legume species (Supplementary Table S2). Manual corrections for start and stop codons and the determination of pseudogenes were performed in Geneious v11 (Biomatters Ltd., Auckland, New Zealand). For the *matK* pseudogene annotated in the cp genome of *C. bonii* (see section “Results”), we further mapped

**TABLE 1** | Characteristics of the 22 complete chloroplast genomes for *Campylotropis*, including 21 newly generated accessions and the previously published accession of *Campylotropis macrocarpa*.

Sample code	Species name	Size (bp)				GC content (%) total (LSC/SSC/IR)	No. of genes (PCGs/tRNA/rRNA)	GenBank accession	Sample location
		Total	Large single-copy region (LSC)	Small single-copy region (SSC)	Inverted repeat (IR)				
xubo1489	<i>Campylotropis albopubescens</i>	149,165	82,871	18,854	23,720	34.84 (32.23/27.94/42.15)	128 (83/37/8)	OM775444	China. Yunnan: Shiping
S867	<i>Campylotropis bonii</i>	153,122	82,869	18,899	25,677	34.98 (32.31/28.02/41.84)	129 (82/39/8)	OM775455	China. Guangxi: Jingxi
XB-DR-C	<i>Campylotropis brevifolia</i>	148,855	82,648	18,805	23,701	34.83 (32.23/27.90/42.13)	128 (83/37/8)	OM775434	China. Yunnan: Derong
xubo1390	<i>Campylotropis capillipes</i>	152,978	82,903	18,701	25,687	34.95 (32.24/28.13/41.81)	130 (83/39/8)	OM775435	China. Yunnan: Binchuan
xubo1445	<i>Campylotropis delavayi</i>	149,088	82,797	18,851	23,720	34.87 (32.28/27.92/42.15)	128 (83/37/8)	OM775436	China. Yunnan: Heqing
xubo1424	<i>Campylotropis grandifolia</i>	149,165	82,871	18,854	23,720	34.84 (32.23/27.94/42.15)	128 (83/37/8)	OM775437	China. Yunnan: Mile
xubo1429	<i>Campylotropis harmsii</i>	149,291	82,992	18,859	23,720	34.86 (32.23/28.01/42.16)	128 (83/37/8)	OM775438	China. Yunnan: Jinhong
xubo1483	<i>Campylotropis henryi</i>	149,153	82,851	18,904	23,699	34.89 (32.30/28.01/42.17)	128 (83/37/8)	OM775440	China. Yunnan: Xiping
xubo1375	<i>Campylotropis howellii</i>	149,312	82,965	18,823	23,762	34.81 (32.17/27.92/42.13)	128 (83/37/8)	OM775439	China. Yunnan: Tengchong
xubo1430	<i>Campylotropis latifolia</i>	149,176	82,881	18,855	23,720	34.84 (32.23/27.93/42.15)	128 (83/37/8)	OM775441	China. Yunnan: Shiping
--	<i>Campylotropis macrocarpa</i>	148,814	82,566	18,808	23,720	34.86 (32.27/27.89/42.14)	128 (83/37/8)	NC_044100	Jin et al., 2019
xubo1425	<i>Campylotropis cytisoides</i> f. <i>parviflora</i>	148,932	82,655	18,846	23,715	34.83 (32.19/27.93/42.15)	128 (83/37/8)	OM775442	China. Yunnan: Jinhong
xubo1426	<i>Campylotropis pinetorum</i> subsp. <i>velutina</i>	149,227	82,933	18,848	23,723	34.86 (32.24/28.02/42.16)	128 (83/37/8)	OM775443	China. Yunnan: Eshan
xubo1447	<i>Campylotropis polyantha</i>	149,191	82,810	18,941	23,720	34.84 (32.25/27.84/42.16)	128 (83/37/8)	OM775447	China. Yunnan: Dali
xubo1427	<i>Campylotropis polyantha</i> var. <i>tomentosa</i>	149,001	82,772	18,801	23,714	34.83 (32.22/27.88/42.15)	128 (83/37/8)	OM775445	China. Sichuan: Shimian
xubo1481	<i>Campylotropis capillipes</i> subsp. <i>prainii</i>	149,092	82,892	18,746	23,727	34.88 (32.25/28.09/42.16)	128 (83/37/8)	OM775446	China. Yunnan: Eshan
xubo1406	<i>Campylotropis teretiracemosa</i>	149,169	82,868	18,863	23,719	34.82 (32.17/28.04/42.16)	128 (83/37/8)	OM775449	China. Sichuan: Yanyuan
xubo1428	<i>Campylotropis thomsonii</i>	148,963	82,676	18,822	23,732	34.85 (32.23/27.94/42.15)	128 (83/37/8)	OM775450	China. Yunnan: Mengla
xubo1393	<i>Campylotropis trigonoclada</i>	149,227	82,957	18,840	23,715	34.83 (32.18/28.05/42.17)	128 (83/37/8)	OM775451	China. Yunnan: Binchuan
xubo1407	<i>Campylotropis wilsonii</i>	149,113	82,771	18,870	23,736	34.85 (32.26/27.90/42.13)	128 (83/37/8)	OM775452	China. Sichuan: Wenchuan
xubo1434	<i>Campylotropis yunnanensis</i> subsp. <i>filipes</i>	149,122	82,822	18,862	23,719	34.84 (32.24/27.90/42.14)	128 (83/37/8)	OM775453	China. Sichuan: Panzhihua
xubo1435	<i>Campylotropis yunnanensis</i>	148,548	82,269	18,841	23,719	34.90 (32.32/27.95/42.14)	128 (83/37/8)	OM775454	China. Yunnan: Yongsheng

raw reads to the assembled sequence of the *matK* gene, and performed Sanger sequencing to validate the accuracy of the assembled sequence. Raw reads were remapped to 400-bp surroundings of the IRb ends to quantify the IR junctions. Genome map of the cp genomes was generated using the online OrganellarGenome DRAW tool (OGDRAW; Lohse et al.,

2013). To compare the contraction and expansion of IRs among cp genomes of *Campylotropis* and closely related genera, we identified and visualized boundaries of LSC, SSC, and IRs of the 25 whole cp genomes (including 22 *Campylotropis* accessions, two *Lespedeza* accessions, and *Kummerowia striata*) using IRscope (Amiryousefi et al., 2018).

## Repeat Sequence Analysis

For 21 newly generated cp genomes and the published accession of *C. macrocarpa*, SSRs were identified using MISA software (Beier et al., 2017) with parameter settings of 11 for mono-, 6 for di-, 5 for tri-, 4 for tetra-, and 3 for penta- and hexa-nucleotide SSRs. For each of the 22 *Campylotropis* cp genomes, forward, reverse, palindrome, and complementary repeat sequences in LSC, IRb, and SSC regions were identified using REPuter program (Kurtz et al., 2001).

## Molecular Marker Identification

The 22 whole cp genomes were firstly aligned using MAFFT v7 (Katoh and Standley, 2013). To identify hypervariable regions that can be used in species identification and phylogenetic studies for *Campylotropis*, nucleotide diversity ( $P_i$ ) values were calculated in sliding windows along the alignment with a window length of 600 bp and step size of 200 bp.  $P_i$  values of each window were calculated using a custom Python script,<sup>1</sup> with the formula referring to the algorithm implemented in pixy (Korunes and Samuk, 2021) to obtain unbiased estimations of nucleotide diversity in the presence of alignment gaps. Adjacent windows with a  $P_i$  value >0.01 and a number of parsimony informative sites >25 were joined together as one single hypervariable region. The number of singleton variable sites, number of parsimony informative sites, and  $P_i$  values were calculated for each hypervariable region using the custom Python script.

## Phylogenetic Analysis

To estimate the cp-genome-based phylogenetic relationships of *Campylotropis* as well as the tribe Desmodieae, we included the whole cp genomes of 22 *Campylotropis* accessions and 15 outgroups (Supplementary Table S2). The phylogenetic analyses were performed using Maximum likelihoods (ML) and Bayesian inference (BI) methods based on both whole cp genomes and shared protein-coding genes (PCGs). For the former dataset, MAFFT v7 was used to obtain the alignment of 37 whole cp genomes. As for the latter dataset, the shared PCGs were extracted and translated into amino acid sequences, and ClustalW2 (Larkin et al., 2007) was used to align the amino acid sequences. The codon alignment of each PCGs was obtained using PAL2NAL (Suyama et al., 2006). The ML trees were inferred using RAxML v8 (Stamatakis, 2014) based on the alignment of 37 whole cp genomes and the concatenated matrix of 72 PCGs. For each RAxML analysis, GTRGAMMA + I was set as the nucleotide substitution model and 1,000 bootstrap replicates were conducted to determine branch support. The BI analyses were performed using MrBayes v3.2 (Ronquist et al., 2012) with the nucleotide substitution model GTR + G + I (lset nst=6 rates=invgamma). For each analysis, the posterior probability was estimated with two independent Markov Chain Monte Carlo (MCMC) chains (10 million generations and sampled every 1,000 generations) with the preliminary 25% of sampled data discarded as burn-in.

<sup>1</sup>[https://github.com/Fengyaa/Campylotropis\\_cp\\_genome](https://github.com/Fengyaa/Campylotropis_cp_genome)

## Analysis of Selective Pressure

To explore the selective pressure of PCGs in *Campylotropis*, the CODEML program implemented in the PAML v4.9 package (Yang, 2007) was used to estimate the rate of non-synonymous ( $d_N$ ) and synonymous ( $d_S$ ) substitutions for PCGs. In general, the ratio of  $d_N/d_S$  ( $\omega$ ) was supposed to equal 1 when under neutral evolution, a larger  $\omega$  indicates higher positive selection pressure, while a smaller ratio of  $\omega$  indicates higher pressure of negative selection.

All the 37 accessions in the above phylogenetic analysis were included, and the resulting phylogenetic tree was used as the input topology for CODEML. The codon-wise alignments of nucleotide sequences, which were used as the input sequences for CODEML, were generated with PAL2NAL (Suyama et al., 2006) guided by the peptide alignments. To determine whether each shared PCG has undergone a different evolutionary force in different lineages, we ran branch-site models with a one-ratio model (null hypothesis;  $\omega_0$ ) in which all branches share the same  $\omega$  and a two-ratio model in which the foreground branches (*Campylotropis* spp.;  $\omega_i$ ) have a different  $\omega$  (alternative hypothesis;  $\omega_b$ ). Likelihood ratio tests with  $\chi^2$  distribution were used to determine whether the alternative hypothesis significantly differ from the null hypothesis (Chi-square test,  $p < 0.05$ ).

## RESULTS

### Characteristics of *Campylotropis* cp Genomes

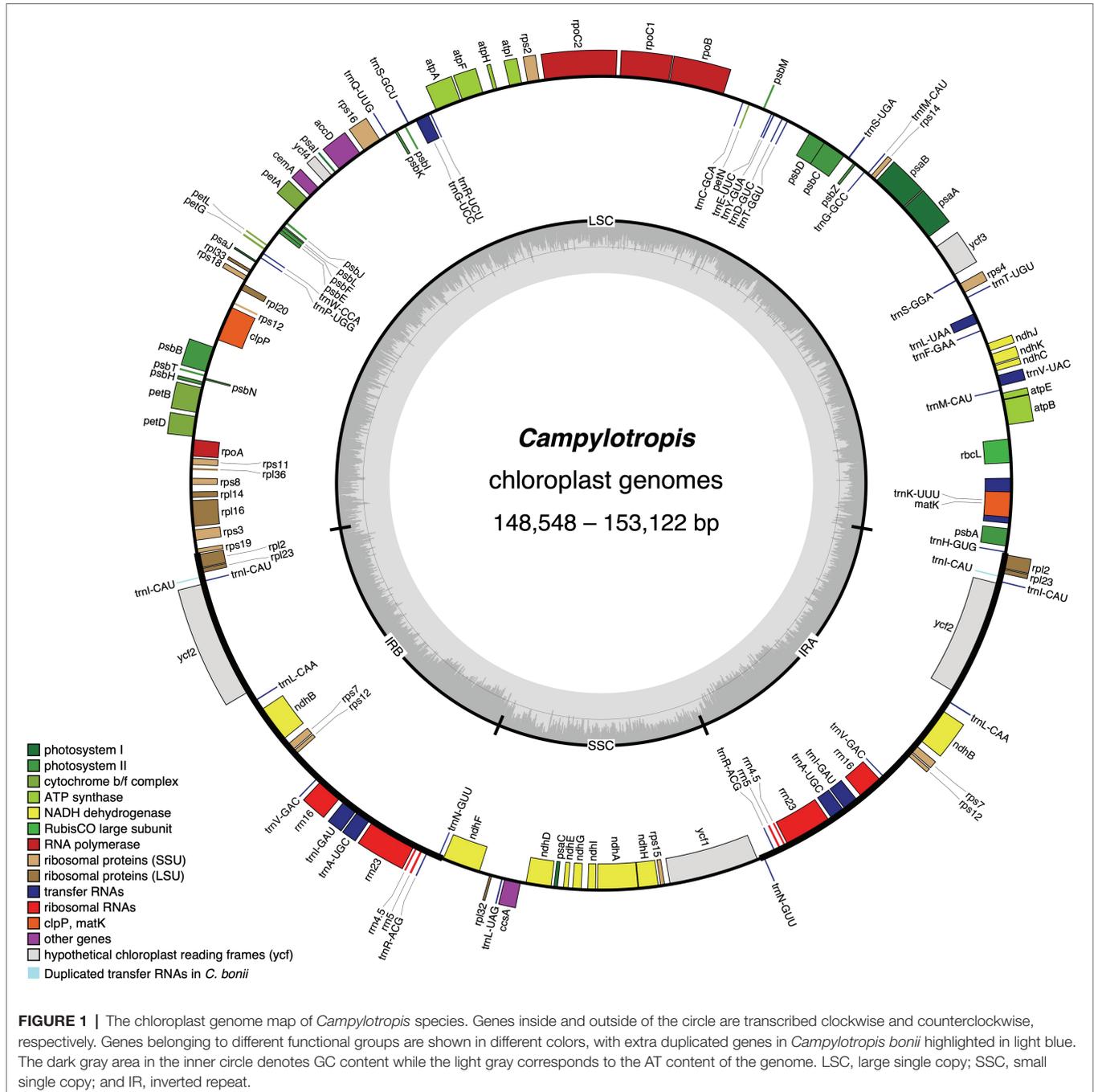
In this study, a total of 21 whole cp genomes of *Campylotropis* were newly generated and were submitted to GenBank under the accession numbers list in Table 1. Taken together with the previously published one of *C. macrocarpa* (NC\_044100), the whole cp genomes of *Campylotropis* ranged from 148,548 bp (*C. yunnanensis*) to 153,122 bp (*C. bonii*), exhibiting a typical quadripartite structure comprising two IR regions (IRa and IRb) of 23,699–25,687 bp, an LSC region of 82,269–82,992 bp, and an SSC region of 18,746–18,941 bp (Table 1). The GC contents of the *Campylotropis* cp genomes were similar (34.81%–34.93%; Table 1). The IRs have the highest GC content (41.81%–42.18%), followed by the LSC region (32.17%–32.32%), and the SSC region (27.84%–28.13%).

The *Campylotropis* cp genomes were similar in gene contents, most of which encode 128 genes, including 83 PCGs, 37 tRNA genes, and eight rRNA genes (all located in the IRs; Table 1; Figure 1). Three species had a few pseudogenes and/or duplicated genes (Table 2). Specifically, *C. capillipes* and *C. bonii* has two more copies of the *trnI-CAU* gene, and *C. bonii* has a pseudogene (*ymatK*; Table 2), which was confirmed by both raw reads mapping and Sanger sequencing (see Supplementary Figure S1 and Supplementary Dataset). Among the 83 PCGs, 77 were unique, and six (*ndhB*, *rpl12*, *rpl23*, *rps7*, *rps12*, and *ycf2*) were duplicated due to their location in the IRs. Likewise, 30 of the tRNA genes are unique, while seven tRNA genes

(*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*) and all four rRNA genes (*rrn23*, *rrn16*, *rrn5*, and *rrn4.5*) were duplicated. Eight PCGs (*petB*, *petD*, *atpF*, *ndhB*, *ndhA*, *rpoC1*, *rpl16*, and *rps16*) and six tRNA genes (*trnA-UGC*, *trnI-GAU*, *trnG-UCC*, *trnL-UAA*, *trnV-UAC*, *trnK-UUU*) contained one intron, while only three PCGs (*rps12*, *ycf3*, and *clpP*) contained two introns (Table 2). In all newly generated *Campylotropis* cp genomes, the 5' end of the *rps12* gene was located in the LSC region, and the 3' end was duplicated in the IRs.

### Comparative Analysis of IR Boundaries

The IR boundary of the assembled cp genomes were quantified by the remapping of short reads, which showed above 300× for the IRb ends and surrounding areas (Supplementary Table S3). We compared the IR boundaries of 25 cp genomes from subtribe Lespedezinae, including *Lespedeza maritima*, *Lespedeza cuneata*, *Kummerowia striata*, and 22 *Campylotropis* accessions, and found a little variation of the expansion/contraction of the IRs (Supplementary Figure S2). The JLA (IRa-LSC) and JSA



**TABLE 2** | Summary of gene contents present in the *Campylotropis* chloroplast genomes.

Group of genes	Name of genes
Ribosomal RNAs	<i>rrn16</i> (x 2), <i>rrn23</i> (x 2), <i>rrn4.5</i> (x 2), <i>rrn5</i> (x 2)
Transfer RNAs	<i>trnA-UGC</i> (1)(x 2), <i>trnI-GAU</i> (1)(x 2), <i>trnL-CAA</i> (x 2), <i>trnN-GUU</i> (x 2), <i>trnR-ACG</i> (x 2), <i>trnV-GAC</i> (x 2), <i>trnI-CAU</i> (x 2)*, <i>trnL-UAG</i> , <i>trnP-UGG</i> , <i>trnW-CCA</i> , <i>trnQ-UUG</i> , <i>trnS-GCU</i> , <i>trnG-UCC</i> (1), <i>trnR-UCU</i> , <i>trnC-GCA</i> , <i>trnE-UUC</i> , <i>trnY-GUA</i> , <i>trnD-GUC</i> , <i>trnT-GGU</i> , <i>trnS-UGA</i> , <i>trnG-GCC</i> , <i>trnM-CAU</i> , <i>trnS-GGA</i> , <i>trnT-UGU</i> , <i>trnL-UAA</i> (1), <i>trnF-GAA</i> , <i>trnV-UAC</i> (1), <i>trnM-CAU</i> , <i>trnK-UUU</i> (1), <i>trnH-GUG</i>
Proteins of small ribosomal subunit	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (x 2), <i>rps8</i> , <i>rps12</i> (2)(x 2), <i>rps14</i> , <i>rps15</i> , <i>rps18</i> , <i>rps19</i> , <i>rps16</i> (1)
Proteins of large ribosomal subunit	<i>rpl2</i> (x 2), <i>rpl14</i> , <i>rpl16</i> (1), <i>rpl20</i> , <i>rpl23</i> (x 2), <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
Subunits of RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> (1), <i>rpoC2</i>
Subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psal</i> , <i>psaJ</i>
Subunits of photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
Subunits of ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> (1), <i>atpH</i> , <i>atpI</i>
Subunits of cytochrome b/f complex	<i>petA</i> , <i>petB</i> (1), <i>petD</i> (1), <i>petG</i> , <i>petL</i> , <i>petN</i>
Subunits of NADH-dehydrogenase	<i>ndhA</i> (1), <i>ndhB</i> (1)(x 2), <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
Large subunit of RuBisCo	<i>rbcL</i>
Acetyl-CoA carboxylase	<i>accD</i>
Cytochrome c biogenesis	<i>ccsA</i>
Envelope membrane protein	<i>cemA</i>
Maturase	<i>matK</i> **
Protease	<i>clpP</i> (2)
Conserved hypothetical chloroplast reading frames	<i>ycf1</i> , <i>ycf2</i> (x 2), <i>ycf3</i> (2), <i>ycf4</i>

(1) Genes with one intron; (2) Genes with two introns; (x 2) Genes with two copies.

\**Campylotropis bonii* and *Campylotropis capillipes* have four copies of *trnI-CAU*.

\*\*The *matK* gene is a pseudogene in *Campylotropis bonii*.

(IRa-SSC) boundaries are highly consistent in the 25 cp genomes, with the former located between *rpl2* and *trnH*, and the latter between *ycf1* and *trnN*. The distances between the JLA boundary and *trnH* were 0–19 bp, while those between the JSA boundary and *ycf1* varied from 128 to 144 bp. The JLB (IRb-LSC) boundaries cut through *rps19* in most species, with 32–48 bp of *rps19* extended into the IRb, while the JLB boundaries of *C. thomsonii* and *C. parviflora* were 88 bp away from *rps19* due to the contraction of IRs. The distance between the JSB (IRb-SSC) boundaries and *ndhF* varied from 2 to 33 bp in most species except *Kummerowia striata*, where *ndhF* extended 11 bp into IRb due to the expansion of IRs.

## Characteristics of Repeat Sequences

The number of SSRs in the *Campylotropis* species varied from 50 in *C. harmsii* to 115 in *C. teretiracemosa* (Supplementary Table S4), in which mononucleotide SSRs were most abundant, followed by component and dinucleotides SSRs (Figure 2A). Among the motifs in the SSRs, A/T, AA/TT, and AT/AT were the most frequently occurring motifs (Figure 2B). Besides, most of the SSRs were located in the LSC (38–56) and SSC (10–18) regions, and very few were

located in the IRs (Supplementary Table S5). REPuter identified 40–71 repeat sequences with length > 30 bp, covering 1,647–4,278 bp in the cp genomes of *Campylotropis* species (Figures 2C,D). Palindromic repeat sequences were most abundant (22–32), followed by forward (14–20) and reverse (2–13) repeat sequences (Figure 2C; Supplementary Table S6). All the repeat sequences with length > 30 bp were located in LSC (33–64) and IRs (6–10), while none of them were identified in the SSC region (Supplementary Table S6). Most of the repeat sequences were less than 100 bp, a few of them were larger than 100 bp (Figure 2D; Supplementary Table S6). Notably, *C. bonii* and *C. capillipes* each had a forward repeat sequence with a length of 2,219 and 2,217 bp, respectively (Figure 2D; Supplementary Table S7). Both repeat sequences were located between *rpl23* and *ycf2* in the IRs, which caused the duplication of *trnI-CAU* and resulted in four copies of this gene (Supplementary Figure S3).

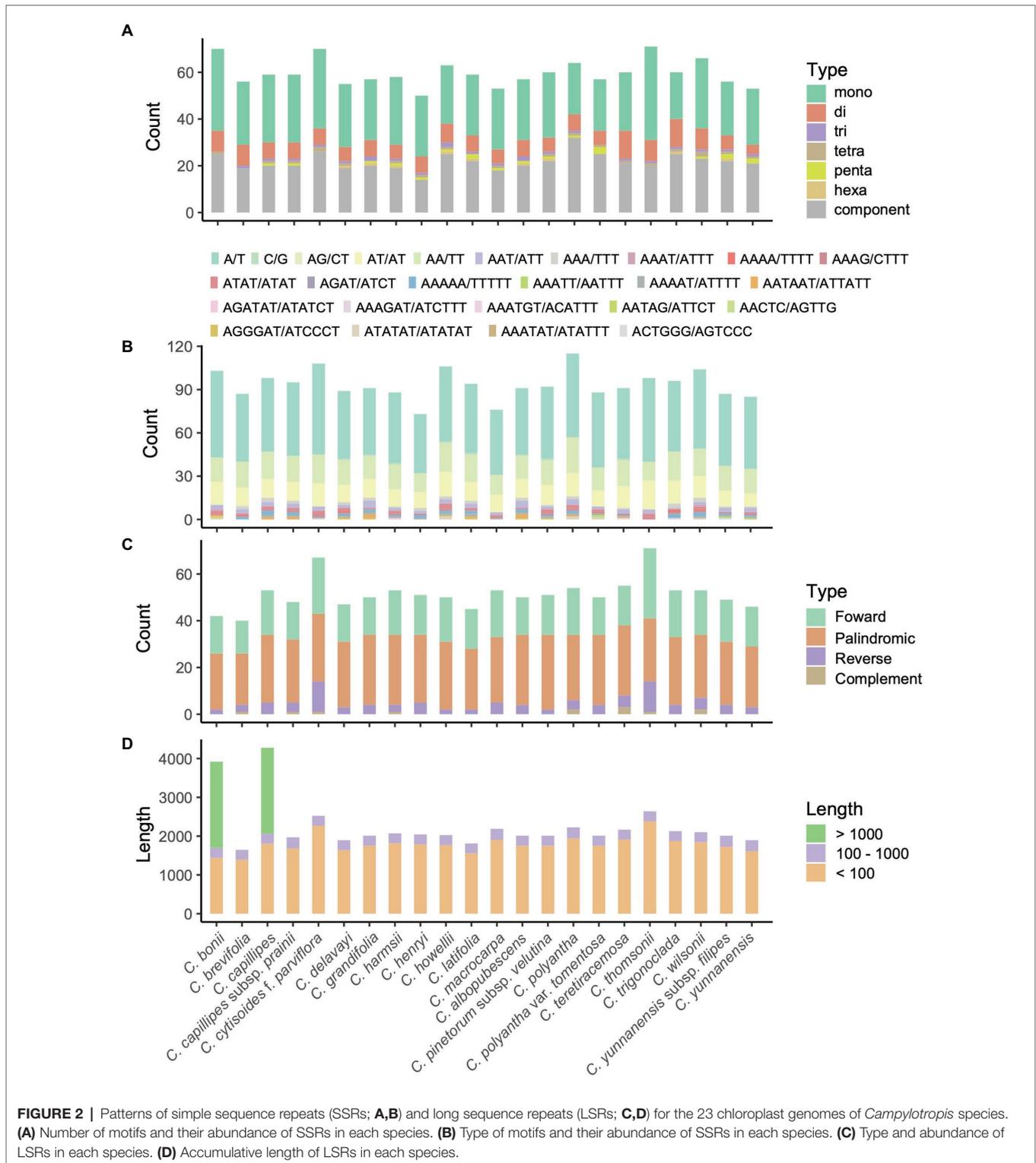
## Identification of Candidate Molecular Markers

Using sliding window analysis, we found that most genetic variations in the cp genomes of *Campylotropis* occurred in the LSC and SSC regions (Figure 3). A total of 13 intergenic spacer regions located in the LSC region, ranging from 547 to 1,995 bp, were identified as potential molecular markers for phylogenetic and population genetic studies (Figure 3; Table 3). Among them, the intergenic spacer of *atpA* and *psbI* (*atpA-psbI*) was the longest (1,995) and contained the greatest number of parsimony informative sites (109), while the intergenic spacer of *ycf4* and *cemA* (*ycf4-cemA*) had the highest Pi value (0.0117).

## Phylogenetic Relationships of *Campylotropis*

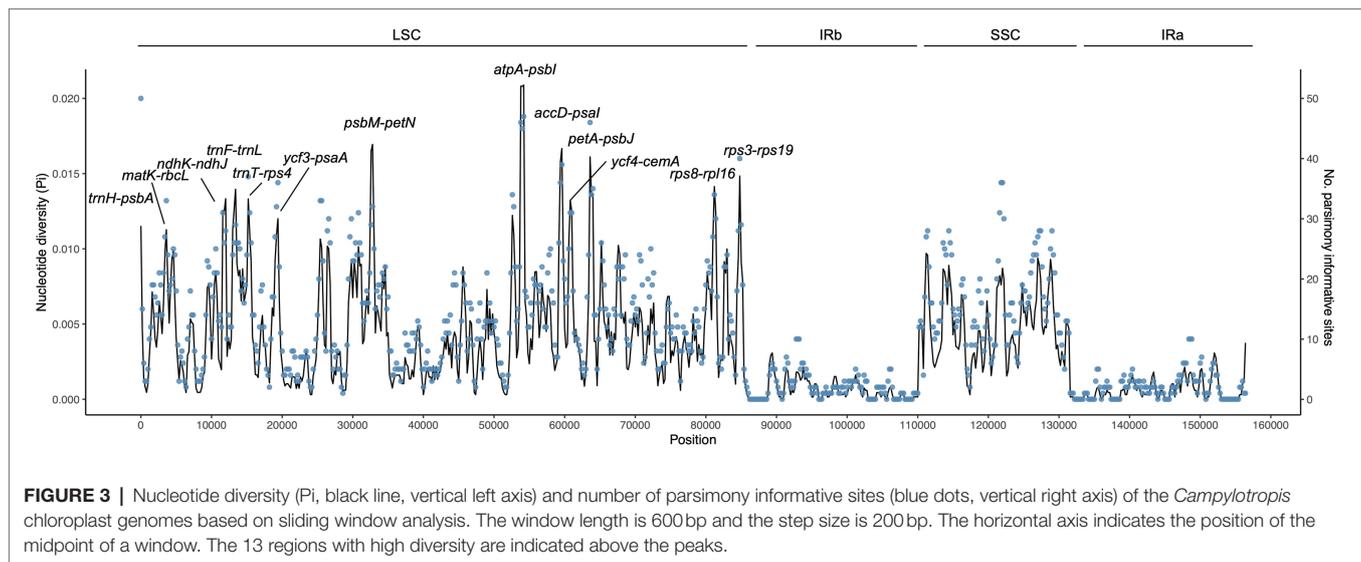
The phylogenetic trees inferred from Maximum likelihood (ML) and Bayesian inference (BI) based on the whole cp genome shared an identical topology and showed little differences in support values (Figure 4). The concatenated alignment of PCGs resulted in similar topologies, with a few differences with regard to the relationships within *Campylotropis* (Supplementary Figures S4, S5). All topologies fully supported the reciprocal monophyly of the two subtribes in tribe Desmodieae [100% bootstrap support (BS) and 1 posterior probability (PP)]. In the subtribe Lespedezinae, *Kummerowia striata* and the two *Lespedeza* species formed a clade (BS=100%, PP=1), and *Campylotropis* was also a monophyletic clade (BS=100%, PP=1).

As for the relationship within *Campylotropis*, both ML and BI trees based on the whole cp genome supported *C. bonii* (lineage A) as sister to the remaining species (Figure 4), and the latter clade (BS=86%, PP=0.99) segregated into two subclades (lineages B and C), each with full support values (BS=100%, PP=1). Lineage B included *C. yunnanensis* subsp. *filipes*, *C. yunnanensis*, *C. polyantha* var. *tomentosa*, *C. macrocarpa*, *C. wilsonii*, *C. polyantha*, *C. brevifolia*, *C. cytisoides* f. *parviflora*, and *C. thomsonii*. And, lineage C included *C. albobubescens*, *C. grandifolia*, *C. latifolia*, *C. delavayi*, *C. capillipes*, *C. capillipes*



subsp. *prainii*, *C. pinetorum* subsp. *velutina*, *C. harmsii*, *C. henryi*, *C. howellii*, *C. trigonoclada*, and *C. teretiracemosa*. The ML three based on the PCGs dataset showed the same topology as that based on the whole cp genome with regard to the relationship among the three subclades of *Campylotropis*, albeit the supporting

values were lower (lineage B: BS=93%; lineage B sister to lineage C: BS=79%; **Supplementary Figure S4**). However, the BI inference based on the PCGs dataset revealed a different topology, in which *C. bonnii* was weakly supported to be a sister clade of lineage B (PP=0.604; **Supplementary Figure S5**).



**FIGURE 3 |** Nucleotide diversity (Pi, black line, vertical left axis) and number of parsimony informative sites (blue dots, vertical right axis) of the *Campylotropis* chloroplast genomes based on sliding window analysis. The window length is 600 bp and the step size is 200 bp. The horizontal axis indicates the position of the midpoint of a window. The 13 regions with high diversity are indicated above the peaks.

**TABLE 3 |** Hypervariable regions identified among the 22 cp genomes of *Campylotropis*.

Start	End	Length	# SVS	# PIP	Pi	Gene name
0	547	547	101	50	0.01154744	<i>trnH-psbA</i>
3,248	4,977	1,729	141	72	0.00815403	<i>matK-rbcL</i>
11,237	12,215	978	61	39	0.00847844	<i>ndhK-ndhJ</i>
12,770	13,801	1,031	241	58	0.01136576	<i>tmF-tmL</i>
14,560	15,451	891	158	49	0.01037634	<i>trnT-rps4</i>
18,421	19,380	959	71	43	0.0078191	<i>ycf3-psaA</i>
31,452	32,596	1,144	116	44	0.01010241	<i>psbM-petN</i>
51,126	53,121	1,995	269	109	0.01122222	<i>atpA-psbI</i>
57,121	58,064	943	84	46	0.01088579	<i>accD-psaI</i>
58,636	59,422	786	48	37	0.01167507	<i>ycf4-cemA</i>
61,206	62,172	966	77	52	0.01052879	<i>petA-psbJ</i>
78,264	79,190	926	147	42	0.00917555	<i>rps8-rpl16</i>
81,692	82,566	874	163	55	0.01090692	<i>rps3-rps19</i>

The start and end positions are referred to *Campylotropis macrocarpa*. # SVS: number of singleton variable sites; # PIP: number of parsimony informative sites.

## Selective Pressure of cp Genes in *Campylotropis*

A total of 68 shared PCGs were subjected to the selective pressure analysis (Supplementary Table S8). Most of the genes were subjected to purifying selection ( $\omega < 1$ ; Figure 5). Using the likelihood ratio test, we found that 11 genes showed significantly different selective pressure in *Campylotropis* (Figure 5; Supplementary Table S8). Among them, two genes (*atpF* and *rps19*) showed obvious signatures of positive selection ( $\omega_t > 1$ ,  $p < 0.05$ ) in *Campylotropis* and eight genes (*ndhC*, *ndhD*, *psbA*, *rpoC1*, *rpoC2*, *rps4*, *ycf1*, and *ycf2*) evolved faster in *Campylotropis* than in the background branches ( $\omega_t > \omega_b$ ,  $p < 0.05$ ; Figure 5; Supplementary Table S8).

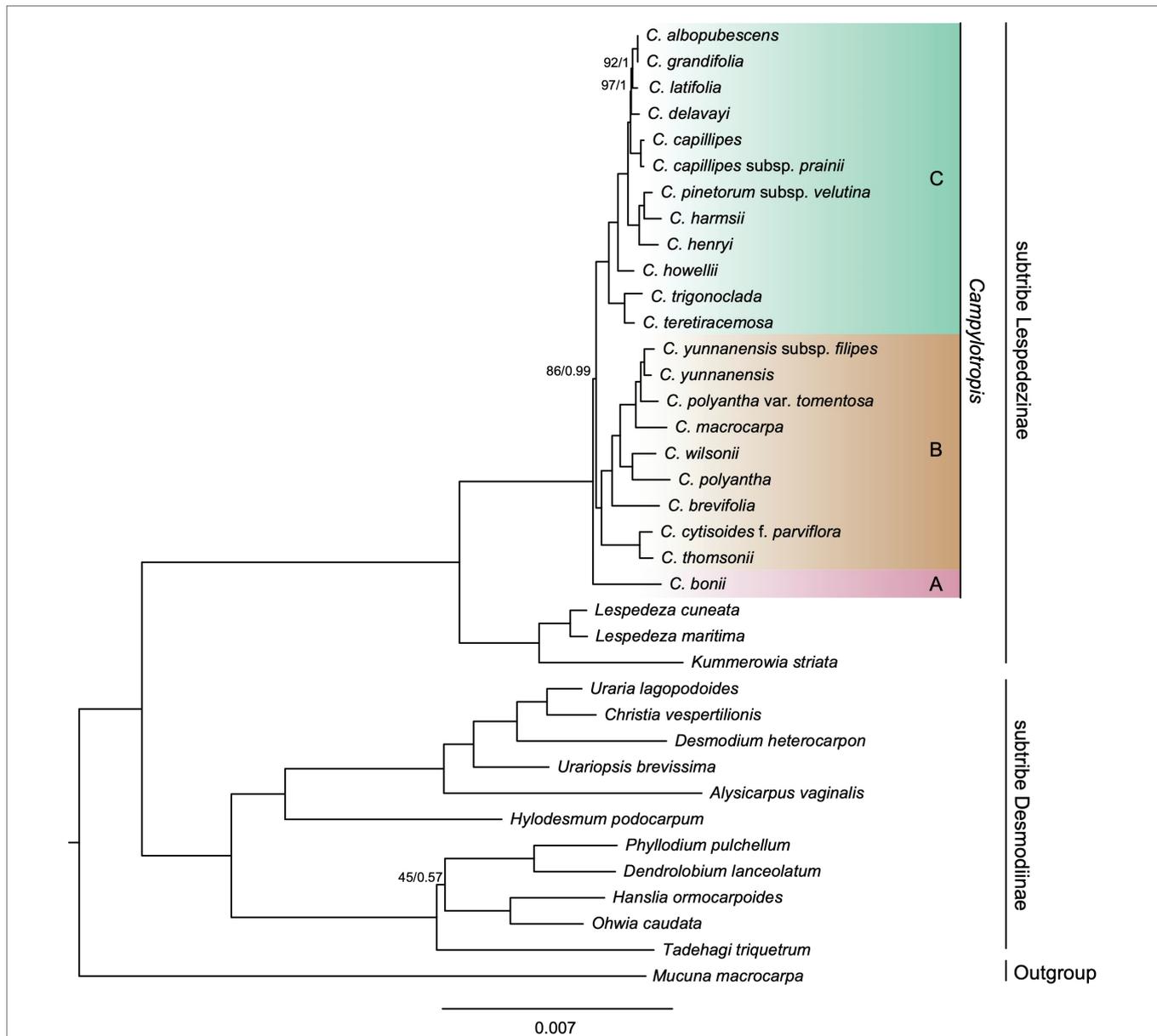
## DISCUSSION

### Variations and Evolution of Whole cp Genomes in *Campylotropis*

The 21 newly assembled and one previously published *Campylotropis* cp genomes showed little variation in genome

structure and genome length, as found in other legume species (Wang et al., 2018; Oyebejani et al., 2020; Zhang et al., 2020; Liao et al., 2021). The *Campylotropis* cp genomes exhibit the typical quadripartite structure and no large structural variant was found (Table 1). The genome length of these species was similar (148,548–153,122 bp) and fell within the range of subfamily Papilionoideae (c. 140–160 kb; Oyebejani et al., 2020). Other genome features, including lengths of LSC, SSC, and IRs, expansion and contraction of IR boundaries, number of genes, GC content, the pattern of SSRs also varied little within this genus, which is comparable to other genera from the legume family (e.g., Oyebejani et al., 2020; Liao et al., 2021).

Despite the general homogeneity characteristics mentioned above, there are some interesting inconsistencies worth mentioning in *Campylotropis* cp genomes. Previous studies demonstrated that expansion and contraction of IRs substantially contribute to the change in the size of cp genomes (Ruhlman and Jansen, 2014; Zheng et al., 2017; Gu et al., 2020). In our study, the JLB (IRb-LSC) boundaries cut through *rps19* in most species, except in *C. thomsonii* and *C. parviflora*, where



**FIGURE 4 |** Phylogenetic tree obtained using the Maximum Likelihood (ML) and Bayesian Inference (BI) method for *Campylotropis* spp. and closely related species based on whole cp genomes. Numbers above branches indicate ML bootstrap supports (BS; before the slash) and Bayesian posterior probabilities (PP; after the slash). The full support values are not indicated.

JLB was located between *rps19* and *rpl2*, causing less than 100-bp length variation of the IRs (Supplementary Figure S2). However, the cp genomes of *C. bonii* and *C. capillipes* were 3–4 kb longer than the rest without showing any significant signal of IR expansion (Figure 2; Supplementary Figure S2). Both cp genomes have a ~2 kb long sequence repeat in each IR region, causing a ~4 kb increase in total genome length. These results indicate that similar to nuclear genomes (Bennetzen et al., 2005), dynamics in repeat sequences rather than expansion and contraction of IRs played an important role in the length variation of *Campylotropis* cp genomes. The long sequence repeats also caused duplication of *trnI-CAU* and resulted in

four copies of this gene (Supplementary Figure S1; Supplementary Table S7).

The *Campylotropis* cp genomes showed moderate sequence variation, most occurring in the LSC region (Figure 3). Consequently, all 13 candidate molecular markers were located in the LSC region, which may be useful in further studies of species delimitation, phylogenetic, and population genetic studies (Table 3). Many of these molecular markers have been reported in other studies, such as *trnH-psbA* (Li et al., 2021), *accD-psaI* (Chen et al., 2021), and *petN-trnD* (Liao et al., 2021). Notably, the *matK* gene, which encodes a protein essential for *in vivo* splicing of Group II introns (Ahlert et al., 2006), is a pseudogene



*Garcinia paucinervis* (Wang et al., 2021). As indicated in Yin et al. (2018), *atpF* gene is highly divergent between deciduous and evergreen sclerophyllous oaks since the former loses its leaves in cold and drought seasons. Despite having  $\omega_r < 1$ , eight genes (*ndhC*, *ndhD*, *psbA*, *rpoC1*, *rpoC2*, *rps4*, *ycf1*, and *ycf2*) significantly accelerated their evolution in *Campylotropis* compared to background branches ( $\omega_r > \omega_b$ ,  $p < 0.05$ ). Some of them were reported to be under significant positive selection in other taxa, such as *ycf1* in seed plants (Zheng et al., 2017), *ndhC* in *Echinacanthus* (Gao et al., 2019), and *rpoC2* in *Rehmannia* (Zeng et al., 2017). Therefore, these positively selected and fast-evolving genes may play an important role in the adaptation of *Campylotropis* species to arid soils and various types of habitats.

## CONCLUSION

In this study, we assembled 21 whole cp genomes for *Campylotropis* spp. Comparative analysis of the cp genome size, structure, expansion and contraction of IR boundaries, number of genes, GC content, and pattern of SSRs revealed high similarities among the *Campylotropis* cp genomes. The activities of long sequence repeats contributed to the variation in genome size and gene content in *Campylotropis* cp genomes. The *Campylotropis* cp genomes showed moderate sequence variation, and 13 candidate regions were identified for further studies of species identification and phylogenetic studies. We also reported one more case of *matK* pseudogene for legume species in *C. bonii*. The phylogenetic analysis confirmed the monophyly of *Campylotropis* and the sister relationship between *Lespedeza* and *Kummerowia*, the latter two genera were then sister to *Campylotropis*. And, its intrageneric relationships based on genomic scale data were firstly reported in this study. The two positively selected genes (*atpF* and *rps19*) and eight fast-evolving genes identified in this study may help us to understand the adaptation of *Campylotropis* species.

## REFERENCES

- Ahlert, D., Piepenburg, K., Kudla, J., and Bock, R. (2006). Evolutionary origin of a plant mitochondrial group II intron from a reverse transcriptase/maturase-encoding ancestor. *J. Plant Res.* 119, 363–371. doi: 10.1007/s10265-006-0284-0
- Amiryousefi, A., Hyvönen, J., and Poczai, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 34, 3030–3031. doi: 10.1093/bioinformatics/bty220
- Barham, J. (1997). 327. *Campylotropis polyantha*: Leguminosae-Papilionoideae. *Curtis's Bot. Mag.* 203–207.
- Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Bennetzen, J. L., Ma, J., and Devos, K. M. (2005). Mechanisms of recent genome size variation in flowering plants. *Ann. Bot.* 95, 127–132. doi: 10.1093/aob/mci008
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Braukmann, T. W. A., Broe, M. B., Stefanović, S., and Freudenstein, J. V. (2017). On the brink: the highly reduced plastomes of nonphotosynthetic Ericaceae. *New Phytol.* 216, 254–266. doi: 10.1111/nph.14681
- Bunge, A. (1835). *Plantarum mongholicorum-chinensium decas prima*. *Uchenyya Zap. Imp. Kazan. Univ.* 4, 154–180.

## DATA AVAILABILITY STATEMENT

The sequences and annotations of the newly generated chloroplast genomes of *Campylotropis* species were deposited in the National Center for Biotechnology Information (NCBI) GenBank database under the accession numbers list in **Table 1**.

## AUTHOR CONTRIBUTIONS

YF, X-FG, and BX conceived and designed the study. BX, H-ND, J-YZ, and ML collected the sample. YF, J-YZ, L-SJ, and XL analyzed the data. YF wrote the manuscript. BX revised the paper. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Natural Science Foundation of China (grant no. 31570196), the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (grant no. 2019QZKK0502), and Wild Plants Sharing and Service Platform of Sichuan Province.

## ACKNOWLEDGMENTS

The authors thank Shi-Qi Li and Xue-Ping Fan for their help with sample collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.895543/full#supplementary-material>

- Cai, J., Ma, P. F., Li, H. T., and Li, D. Z. (2015). Complete plastid genome sequencing of four *Tilia* species (Malvaceae): a comparative analysis and phylogenetic implications. *PLoS One* 10:e0142705. doi: 10.1371/journal.pone.0142705
- Chen, J., Zang, Y., Shang, S., Liang, S., Zhu, M., Wang, Y., et al. (2021). Comparative chloroplast genomes of Zosteraceae species provide adaptive evolution insights into seagrass. *Front. Plant Sci.* 12:741152. doi: 10.3389/fpls.2021.741152
- Cvijović, I., Good, B. H., and Desai, M. M. (2018). The effect of strong purifying selection on genetic diversity. *Genetics* 209, 1235–1278. doi: 10.1534/genetics.118.301058
- Daniell, H., Lin, C.-S., Yu, M., and Chang, W.-J. (2016). Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol.* 17:134. doi: 10.1186/s13059-016-1004-2
- Dopp, I. J., Yang, X., and Mackenzie, S. A. (2021). A new take on organelle-mediated stress sensing in plants. *New Phytol.* 230, 2148–2153. doi: 10.1111/nph.17333
- Du, Q., Bi, G., Mao, Y., and Sui, Z. (2016). The complete chloroplast genome of *Gracilaria lemaneiformis* (Rhodophyta) gives new insight into the evolution of family Gracilariaceae. *J. Phycol.* 52, 441–450. doi: 10.1111/jpy.12406
- Fu, P. Y. (1987). A study of the genus *Campylotropis* Bunge in China. *Bull. Bot. Res.* 7, 11–55.
- Gao, C., Deng, Y., and Wang, J. (2019). The complete chloroplast genomes of *Echinacanthus* species (Acanthaceae): phylogenetic relationships, adaptive

- evolution, and screening of molecular markers. *Front. Plant Sci.* 9:1989. doi: 10.3389/fpls.2018.01989
- Gao, X., Zhang, X., Meng, H., Li, J., Zhang, D., and Liu, C. (2018). Comparative chloroplast genomes of *Paris* sect. *Marmorata*: insights into repeat regions and evolutionary implications. *BMC Genomics* 19, 133–144. doi: 10.1186/s12864-018-5281-x
- Gu, L., Su, T., An, M. T., and Hu, G. X. (2020). The complete chloroplast genome of the vulnerable *Oreocharis esquirolii* (Gesneriaceae): structural features, comparative and phylogenetic analysis. *Plan. Theory* 9, 1–16. doi: 10.3390/plants9121692
- Han, B., Jiang, P., Liu, W., Xu, H., Li, Y., Li, Z., et al. (2018). Role of daucosterol linoleate on breast cancer: studies on apoptosis and metastasis. *J. Agric. Food Chem.* 66, 6031–6041. doi: 10.1021/acs.jafc.8b01387
- Howe, C. J., Barbrook, A. C., Koumandou, V. L., Nisbet, R. E. R., Symington, H. A., Wightman, T. F., et al. (2003). Evolution of the chloroplast genome. *Philos. Trans. R. Soc. B Biol. Sci.* 358, 99–107. doi: 10.1098/rstb.2002.1176
- Huang, P., Ohashi, H., Iokawa, Y., and Nemoto, T. (2010). “Fabaceae, tribe desmodieae,” in *Flora of China*. eds. C. Y. Wu, P. H. Raven and D. Y. Hong (St. Louis: Science Press, Beijing, and Missouri Botanical Garden), 262–311.
- Iokawa, Y., and Ohashi, H. (2008). *Campylotropis* (Leguminosae) of China, an enumeration and distribution. *J. Japanese Bot.* 83, 36–59.
- Jabbour, F., Gaudeul, M., Lambourdière, J., Ramstein, G., Hassanin, A., Labat, J. N., et al. (2018). Phylogeny, biogeography and character evolution in the tribe Desmodieae (Fabaceae: Papilionoideae), with special emphasis on the new Caledonian endemic genera. *Mol. Phylogenet. Evol.* 118, 108–121. doi: 10.1016/j.ympev.2017.09.017
- Jin, D. P., Choi, I. S., and Choi, B. H. (2019). Plastid genome evolution in tribe Desmodieae (Fabaceae: Papilionoideae). *PLoS One* 14, 1–17. doi: 10.1371/journal.pone.0218743
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., Depamphilis, C. W., Yi, T. S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21, 1–31. doi: 10.1186/s13059-020-02154-5
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Keller, J., Rousseau-Guettin, M., Martin, G. E., Morice, J., Boutte, J., Coissac, E., et al. (2017). The evolutionary fate of the chloroplast and nuclear rps16 genes as revealed through the sequencing and comparative analyses of four novel legume chloroplast genomes from *Lupinus*. *DNA Res.* 24, 343–358. doi: 10.1093/dnares/dsx006
- Kocyan, A., Vogel, E. F. D., Conti, E., and Gravendeel, B. (2008). Molecular phylogeny of *Aerides* (Orchidaceae) based on one nuclear and two plastid markers: A step forward in understanding the evolution of the Aeridinae. *Mol. Phylogenet. Evol.* 48, 422–443. doi: 10.1016/j.ympev.2008.02.017
- Korunes, K. L., and Samuk, K. (2021). Pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol. Ecol. Resour.* 21, 1359–1368. doi: 10.1111/1755-0998.13326
- Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Li, Y., Dong, Y., Liu, Y., Yu, X., Yang, M., and Huang, Y. (2021). Comparative analyses of *euonymus* chloroplast genomes: genetic structure, screening for loci with suitable polymorphism, positive selection genes, and phylogenetic relationships within Celastrineae. *Front. Plant Sci.* 11:593984. doi: 10.3389/fpls.2020.593984
- Li, C., Zhao, Y., Xu, Z., Yang, G., Peng, J., and Peng, X. (2020). Initial characterization of the chloroplast genome of *Vicia sepium*, an important wild resource plant, and related inferences about its evolution. *Front. Genet.* 11:73. doi: 10.3389/fgenet.2020.00073
- Liao, M., Gao, X.-F., Zhang, J.-Y., Deng, H.-N., and Xu, B. (2021). Comparative chloroplast genomics of *Sophora* species: evolution and phylogenetic relationships in the early-diverging legume subfamily Papilionoideae (Fabaceae). *Front. Plant Sci.* 12:778933. doi: 10.3389/fpls.2021.778933
- Lohse, M., Drechsel, O., Kahlau, S., and Bock, R. (2013). OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41, W575–W581. doi: 10.1093/nar/gkt289
- Lu, R. S., Li, P., and Qiu, Y. X. (2017). The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: comparative genomic and phylogenetic analyses. *Front. Plant Sci.* 7:2054. doi: 10.3389/fpls.2016.02054
- Luo, Y., Ma, P. F., Li, H. T., Yang, J. B., Wang, H., and Li, D. Z. (2016). Plastid phylogenomic analyses resolve Tofieldiaceae as the root of the early diverging monocot order Alismatales. *Genome Biol. Evol.* 8, 932–945. doi: 10.1093/gbe/evv260
- Moseley, R. C., Mewalal, R., Motta, F., Tuskan, G. A., Haase, S., and Yang, X. (2018). Conservation and diversification of circadian rhythmicity between a model crassulacean acid metabolism plant *kalanchoë fedtschenkoi* and a model C3 photosynthesis plant *arabidopsis thaliana*. *Front. Plant Sci.* 9:1757. doi: 10.3389/fpls.2018.01757
- Oyebanji, O., Zhang, R., Chen, S. Y., and Yi, T. S. (2020). New insights into the plastome evolution of the Millettioideae/Phaseoloideae clade (Papilionoideae, Leguminosae). *Front. Plant Sci.* 11:151. doi: 10.3389/fpls.2020.00151
- Patwardhan, A., Ray, S., and Roy, A. (2014). Molecular markers in phylogenetic studies—a review. *J. Phylogenetics Evol. Biol.* 2:131. doi: 10.4172/2329-9002.1000131
- Qu, X. J., Moore, M. J., Li, D. Z., and Yi, T. S. (2019). PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* 15, 1–12. doi: 10.1186/s13007-019-0435-7
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029
- Ruhlman, T. A., and Jansen, R. K. (2014). The plastid genomes of flowering plants. *Methods Mol. Biol.* 1132, 3–38. doi: 10.1007/978-1-62703-995-6\_1
- Ruhsam, M., Rai, H. S., Mathews, S., Ross, T. G., Graham, S. W., Raubeson, L. A., et al. (2015). Does complete plastid genome sequencing improve species discrimination and phylogenetic resolution in araucaria? *Mol. Ecol. Resour.* 15, 1067–1078. doi: 10.1111/1755-0998.12375
- Sabir, J., Schwarz, E., Ellison, N., Zhang, J., Baeshen, N. A., Mutwakil, M., et al. (2014). Evolutionary and biotechnology implications of plastid genome variation in the inverted-repeat-lacking clade of legumes. *Plant Biotechnol. J.* 12, 743–754. doi: 10.1111/pbi.12179
- Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., et al. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92, 142–166. doi: 10.3732/ajb.92.1.142
- Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Sugiura, M. (1992). The chloroplast genome. *Plant Mol. Biol.* 19, 149–168. doi: 10.1007/BF00015612
- Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612. doi: 10.1093/nar/gkl315
- Wang, Y. H., Wicke, S., Wang, H., Jin, J. J., Chen, S. Y., Zhang, S. D., et al. (2018). Plastid genome evolution in the early-diverging legume subfamily Cercidoideae (Fabaceae). *Front. Plant Sci.* 9:138. doi: 10.3389/fpls.2018.00138
- Wang, Y., Zhao, B., Lu, Z., Shi, Y., and Li, J. (2021). The complete chloroplast genome provides insight into the polymorphism and adaptive evolution of *Garcinia paucinervis*. *Biotechnol. Equip.* 35, 377–391. doi: 10.1080/13102818.2021.1879676
- Wanke, S., Jaramillo, M. A., Borsch, T., Samain, M. S., Quandt, D., and Neinhuis, C. (2007). Evolution of Piperales—*matK* gene and *trnK* intron sequence data reveal lineage specific resolution contrast. *Mol. Phylogenet. Evol.* 42, 477–497. doi: 10.1016/j.ympev.2006.07.007
- Wen, P., Han, H., Wang, R., Wang, N., and Yao, X. (2007). C-glycosylflavones and aromatic glycosides from *Campylotropis hirtella* (Franch.) Schindl. *Asian J. Tradit. Med.* 2, 149–153.
- Wicke, S., Schneeweiss, G. M., dePamphilis, C. W., Müller, K. F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol. Biol.* 76, 273–297. doi: 10.1007/s11103-011-9762-4

- Xu, B., Wu, N., Gao, X. F., and Zhang, L. B. (2012). Analysis of DNA sequences of six chloroplast and nuclear genes suggests incongruence, introgression, and incomplete lineage sorting in the evolution of *lespedeza* (Fabaceae). *Mol. Phylogenet. Evol.* 62, 346–358. doi: 10.1016/j.ympev.2011.10.007
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Yin, K., Zhang, Y., Li, Y., and Du, F. K. (2018). 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.* 19:1042. doi: 10.3390/ijms19041042
- Zeng, S., Zhou, T., Han, K., Yang, Y., Zhao, J., and Liu, Z. L. (2017). The complete chloroplast genome sequences of six *Rehmannia* species. *Genes* 8:103. doi: 10.3390/genes8030103
- Zhang, S. D., Jin, J. J., Chen, S. Y., Chase, M. W., Soltis, D. E., Li, H. T., et al. (2017). Diversification of Rosaceae since the late cretaceous based on plastid phylogenomics. *New Phytol.* 214, 1355–1367. doi: 10.1111/nph.14461
- Zhang, R., Wang, Y. H., Jin, J. J., Stull, G. W., Bruneau, A., Cardoso, D., et al. (2020). Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Syst. Biol.* 69, 613–622. doi: 10.1093/sysbio/syaa013
- Zhang, L., Wang, S., Su, C., Harris, A. J., Zhao, L., Su, N., et al. (2021). Comparative chloroplast genomics and phylogenetic analysis of *Zygophyllum* (Zygophyllaceae) of China. *Front. Plant Sci.* 12:723622. doi: 10.3389/fpls.2021.723622
- Zhao, C., Wang, Y., Chan, K. X., Marchant, D. B., Franks, P. J., Randall, D., et al. (2019). Evolution of chloroplast retrograde signaling facilitates green plant adaptation to land. *Proc. Natl. Acad. Sci. U. S. A.* 116, 5015–5020. doi: 10.1073/pnas.1812092116
- Zheng, X.-M., Wang, J., Feng, L., Liu, S., Pang, H., Qi, L., et al. (2017). Inferring the evolutionary mechanism of the chloroplast genome size by comparing whole-chloroplast genome sequences in seed plants. *Sci. Rep.* 7, 1–10. doi: 10.1038/s41598-017-01518-5

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Feng, Gao, Zhang, Jiang, Li, Deng, Liao and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.