RESEARCH

The complete plastomes of seven *Peucedanum* plants: comparative and phylogenetic analyses for the *Peucedanum* genus

Chang-Kun Liu, Jia-Qing Lei, Qiu-Ping Jiang, Song-Dong Zhou[*] and Xing-Jin He[*]

Abstract

Background: The *Peucedanum* genus is the backbone member of Apiaceae, with many economically and medically important plants. Although the previous studies on *Peucedanum* provide us with a good research basis, there are still unclear phylogenetic relationships and many taxonomic problems in *Peucedanum*, and a robust phylogenetic framework of this genus still has not been obtained, which severely hampers the improvement and revision of taxonomic system for this genus. The plastid genomes possessing more variable characters have potential for reconstructing a robust phylogeny in plants.

Results: In the current study, we newly sequenced and assembled seven *Peucedanum* plastid genomes. Together with five previously published plastid genomes of *Peucedanum*, we performed a comprehensively comparative analyses for this genus. Twelve *Peucedanum* plastomes were similar in terms of genome structure, codon bias, RNA editing sites, and SSRs, but varied in genome size, gene content and arrangement, and border of SC/IR. Fifteen mutation hotspot regions were identified among plastid genomes that can serve as candidate DNA barcodes for species identification in *Peucedanum*. Our phylogenetic analyses based on plastid genomes generated a phylogeny with high supports and resolutions for *Peucedanum* that robustly supported the non-monophyly of genus *Peucedanum*.

Conclusion: The plastid genomes of *Peucedanum* showed both conservation and diversity. The plastid genome data were efficient and powerful for improving the supports and resolutions of phylogeny for the complex *Peucedanum* genus. In summary, our study provides new sights into the plastid genome evolution, taxonomy, and phylogeny for *Peucedanum* species.

Keywords: Apiaceae, *Peucedanum*, Plastome evolution, Phylogenomics, DNA barcoding

Background

Peucedanum L. is one of the largest genera of Apiaceae [1–3], which was once placed in the tribe Peucedaneae [1, 4, 5], but now in the tribe Selineae [2]. The genus comprises 100-120 species worldwide that are widely distributed in Eurasia and South Africa (and sometimes Australia) [2, 3, 6], with Europe and East Asia as distribution centers [7]. Of those, forty species are distributed in China with 33 of them endemic [3].

The genus *Peucedanum* is taxonomically notorious within Apiaceae family, especially described as "*Peucedanum* problem" by Downie et al. [8]. Its members are characterized by dorsally compressed mericarps with slightly prominent dorsal ribs, narrowly winged lateral ribs, as well as a broad commissure [2, 3]. However, the

*Correspondence: zsd@scu.edu.cn; xjhe@scu.edu.cn
Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 2 of 14

genus is extremely heterogenous and exhibits great diversity in life-forms, leaf and fruit structures, and chemical constituents [9]. Hence, several researchers are prone to divide this genus into smaller and presumably more natural units. For example, Pimenov and Leonov [5] suggested that all members of *Peucedanum* except 8-10 species included in sect. *Peucedanum* should be transferred to other genera. Based on morphological and phytochemical evidences, Reduron et al. [10] separated the genera *Cervaria* Wolf, *Imperatoria* L., *Oreoselinum* Mill., *Pteroselinum* Rchb., *Thysselinum* Adans., *Xanthoselinum* Schur and *Holandrea* Reduron from *Peucedanum*. Winter et al. [11] established three new genera (*Afrosciadium* P.J.D. Winter, *Nanobubon* Magee and *Notobubon* B.-E. van Wyk) to accommodate the African peucedanoid species and transferred 24 *Peucedanum* species into *Afroligusticum* C. Norman, *Cynorhiza* Eckl. & Zeyh., and *Lefebvrea* A. Rich. However, due to the varied morphological features of leaf division, bracteoles, and mericarps, distinguishing separate genera from *Peucedanum* is extremely difficult [2, 3]. Therefore, the generic limits of *Peucedanum* based on morphological characters faces challenges.

A robust phylogenetic framework could provide a valuable information to aid the generic delimitation of *Peucedanum*. Previously, a few molecular phylogenies of *Peucedanum* based on single or multiple-locus DNA sequence data, such as nuclear ribosomal DNA internal transcribed spacer (ITS), plastid DNA *rpl*16 and *rps*16 intron, have been performed, yet these analyses failed to recognize *Peucedanum* as a monophyletic group [2, 12–16]. This phenomenon infers that re-evaluating the generic limits of *Peucedanum* may be essential. Nevertheless, weak supports and low resolutions of these phylogenetic trees could not provide sufficient information to support the improvement of taxonomy for *Peucedanum*. Therefore, additional molecular data are urgent to reconstruct a strong phylogeny.

In addition, several species of *Peucedanum* are highly appreciated as traditional medicinal herbs due to their versatile therapeutic properties [17]. Among them, *Peucedanum praeruptorum* Dunn, known as "Baihu Qianhu", is an excellent representation. The dried root of *P. praeruptorum* has been utilized as traditional Chinese medicine for more than 1500 years, which is generally used to treat respiratory diseases, pulmonary hypertension, chest pain, as well as symptomatic coughs and dyspnea [18]. However, most *Peucedanum* species exhibit abundant intraspecific variations in morphology that make it difficult to accurately identify species. In order to assure medicinal quality, it is, therefore, necessary to develop specific DNA marker for *Peucedanum* species authentication.

The plastid genome (plastome) is one of the three DNA genomes (with nuclear and mitochondrial genomes) in plants. The genome is uniparentally inherited, lacks recombination, and possesses highly variable characters in flowering plants; hence, it has the potential to significantly improve the supports and resolutions of the phylogeny [19–22]. Furthermore, a typical plastome comprises two inverted repeats regions (IRs) of 22-25 kb separated by the large single copy region (LSC) of 82-90 kb and small single copy region (SSC) of 15-20 kb and generally encodes 110-130 unique genes [23, 24]. Comparative analysis of plastome could reveal the diversity of plastome in structural organization, gene arrangement and content that deepens our understanding of adaptive evolution for plant lineages and identify suitable mutation hotspots for species authentication [21, 25, 26]. Hence, with the development of next-generation sequencing and bioinformatics technologies, plastomes have been extensively and successfully used for plant phylogenetic analyses and development of specific DNA barcodes in recent years [25–32].

Currently, although six plastomes of *Peucedanum* species were submitted in GenBank [33–36], the plastid phylogenomic analysis of the genus has not been conducted. In this study, we newly sequenced the plastomes of seven *Peucedanum* taxa. In conjunction with the previously reported five plastomes of *Peucedanum*, we carried out a comprehensive analysis of plastomes for this taxonomically difficult plant group. Our aims were to: (1) investigate the plastome features of *Peucedanum* plants; (2) screen out suitable mutation hotspot regions from plastome as candidate DNA barcodes for species identification of *Peucedanum*; (3) test the power of plastome for improving the supports and resolutions of phylogeny in the complex *Peucedanum* genus. Overall, our results will well lay the foundation for the phylogenetic and taxonomic studies of *Peucedanum*.

Results

Plastome features of *Peucedanum*

Illumina sequencing generated 36,875,778-44,140,972 paired-end clean reads for the seven *Peucedanum* samples. Among them, 712,889 to 6,125,929 reads were mapped to the final assembly. Based on these data, we obtained seven high-quality *Peucedanum* plastomes, with coverage ranging from 730.073× to 6,266.178× (Table S1).

Overall size of plastomes ranged from 142,494 bp (*P. angelicoides* Wolff ex Kretschm.) to 156,899 bp (*P. insolens* Kitag.) for the twelve *Peucedanum* samples (Table 1). All of them shown typically quadripartite structure, including a pair of inverted repeats regions (IRs, 12,594-27,495 bp), a large single copy region

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 3 of 14

Table 1 Comparison of plastome features among *Peucedaum* plants

Taxon	Total length (bp)	LSC (bp)	SSC (bp)	IR (bp)	Total GC content (%)	Total genes (unique)	Protein coding genes (unique)	rRNA genes (unique)	tRNA genes (unique)
P. ampliatum	147,403	92,526	17,519	18,679	37.6	114	80	4	30
P. angelicoides	142,494	99,934	17,372	12,594	37.4	114	80	4	30
P. chujaense	147,839	93,335	17,590	18,457	37.4	114	80	4	30
P. delavayi	155,552	85,276	17,394	26,441	37.6	113	79	4	30
P. harry-smithii var. *grande*	147,046	92,135	17,627	18,642	37.6	113	80	4	29
P. insolens	156,899	84,492	17,417	27,495	37.7	113	79	4	30
P. japonicum	147,592	92,804	17,576	18,606	37.5	114	80	4	30
P. longshengense	147,967	93,265	17,572	18,565	37.5	114	80	4	30
P. mashanense	154,230	86,957	16,665	25,304	37.4	114	80	4	30
P. medicum	152,288	86,645	17,571	24,036	37.5	114	80	4	30
P. praeruptorum	147,197	92,161	17,610	18,713	37.6	113	80	4	29
P. terebinthaceum	147,925	93,368	17,571	18,493	37.5	114	80	4	30

(LSC, 84,492-99,934 bp), and a small single copy region (SSC, 16,665-17,627 bp) (Fig. 1, Table 1). The total GC content of the twelve plastomes ranged from 37.4% to 37.7% (Table 1). The twelve plastomes encoded 113-114 unique genes, including 79-80 protein-coding genes, 29-30 tRNA genes, and four rRNA genes (Table 1, Table S2). The *ycf*15 gene was lost in *P. delavayi* Franch. and *P. insolens*; the *trn*T-GGU gene was absent in *P. praeruptorum* and *P. harry-smithii* var. *grande* (K.T.Fu) Shan et Sheh (Table S2).

In order to analyze the codon usage of *Peucedanum* plastomes, 79 protein-coding genes were extracted and connected for each plastome. These sequences were 66,552-68,130 bp in length and encoded 22,184-22,710 codons. The Leu was encoded by the highest number of codons (2,347-2,404), while the Cys was the least (234-243) in all plastomes (Table S3). In addition, relative synonymous codon usage (RSCU) values of all codons ranged from 0.32 to 2.01 in the twelve plastomes (Table S3). Specifically, RSCU values of 30 codons were greater than 1.00 in all plastomes, whereas the codon AUA with RSCU > 1.00 was only detected in *P. insolens* plastomes (Fig. 2). All codons with RSCU > 1.00 were ended with A/U, except UUG (Fig. 2).

The potential RNA editing sites for 35 protein-coding genes of the twelve plastomes were detected. A total of 56-60 potential RNA editing sites were identified (Table S4, Fig. S1). All detected RNA editing sites were Cytosine to Uracil (C-U) conversion and most of them occurred in the second codon position (42-45), followed by the first codon position (12-16), but no sites situated in the third codon position (Fig. S1A). Moreover, the *ndh*B gene contained the highest number of RNA editing sites ranging from 10 to 11 (Fig. S1B).

The total number of SSRs ranged from 58 to 89 among the twelve *Peucedanum* plastomes (Fig. 3, Table S5). Most of the SSRs distributed in the LSC region for all plastomes (Fig. 3A). Among these SSRs, the mononucleotide repeats were the most abundant (28-54), followed by the dinucleotides (14-21) (Fig. 3B). In addition, bases A and T were the dominant elements for all identified SSRs in the twelve plastomes.

Plastome comparison and hotspots identification

The borders of IRa/SSC, IRb/SSC, and IRb/LSC among the twelve *Peucedanum* plastomes were slightly conserved: the IRa/SSC junctions of most samples were located between *ycf*1 gene and *ndh*F gene, but expanded into *ndh*F gene in *P. delavayi* and *P. angelicoides*; the boundaries of IRb/SSC fell into *ycf*1 gene; the IRb/LSC borders of most samples were located between genes of *trn*L and *trn*H, but extremely expanded into *psb*A gene in *P. angelicoides* (Fig. 4). However, the junctions of IRa/LSC of plastomes within *Peucedanum* genus were divergent and could be classified into four different types. The junctions of IRa/LSC fell into the *rps*19 gene in *P. delavayi* and *P. insolens*, belonging to the type I; the IRa/LSC borders contracted to the intergenic region of *trn*L-*trn*H in *P. angelicoides* (type II) while moved to the intergenic regions of *rpl*2-*trn*I in *P. mashanense* Shan et Sheh and *P. medicum* Dunn (type III); the IRa/LSC borders of most remainder *Peucedanum* plants fell into the *ycf*2 gene, but contracted to the intergenic regions of *ycf*2-*trn*L in *P. chujaense* K. Kim, S.H. Oh, C.S. Kim & C.W. Park and *P. terebinthaceum* (Fisch.) Fisch. ex Turcz. (type IV) (Fig. 4).

The genes arrangement of the twelve *Peucedanum* plastomes was relatively conserved, except for an inversion of the *trn*Y-*trn*D-*trn*E gene detected in *P. japonicum*

Liu *et al. BMC Plant Biology* (2022) 22:101

Page 4 of 14
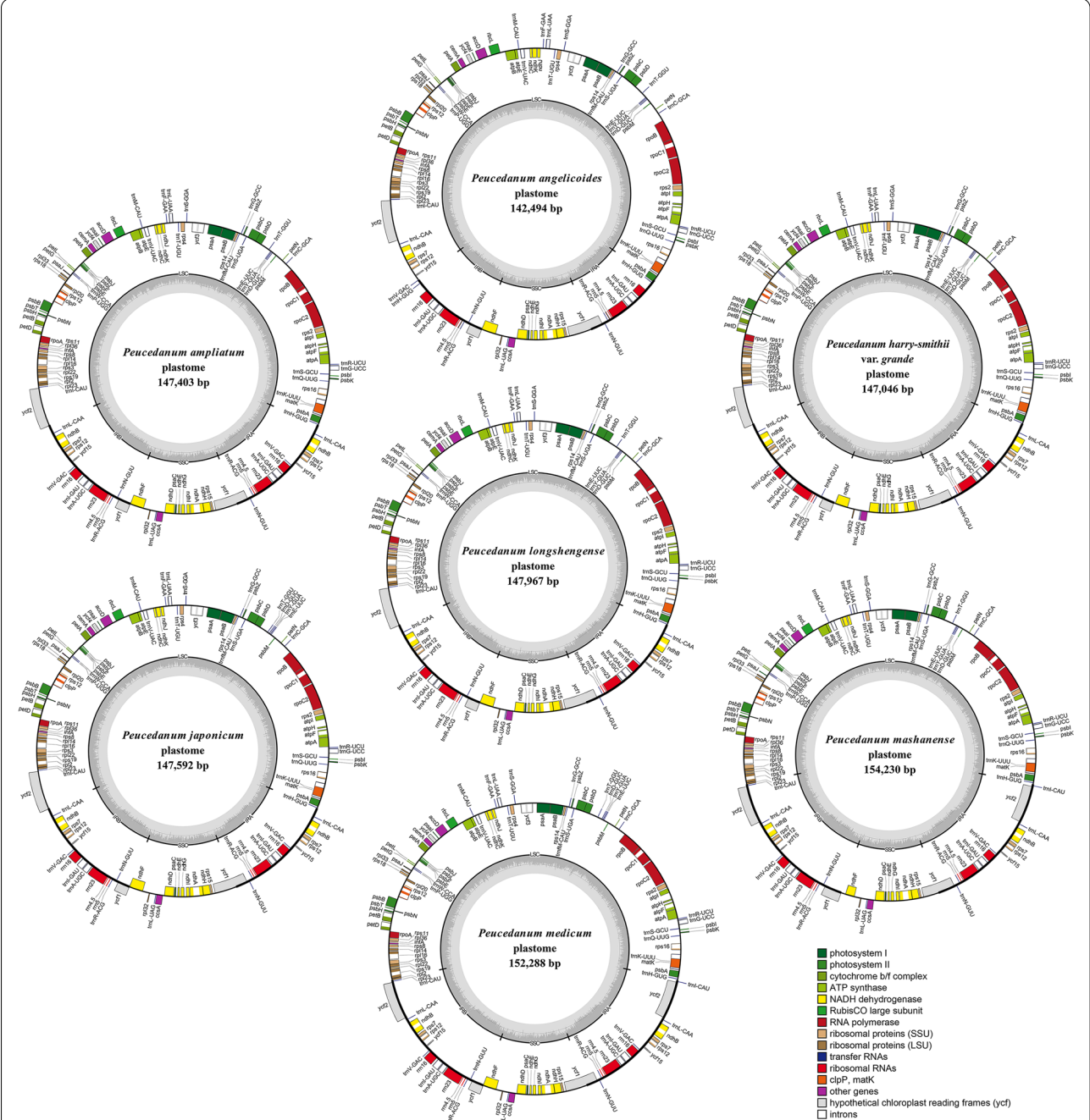
Fig. 1 Maps of seven *Peucedanum* plastomes. Genes shown outside of the outer layer circle are transcribed clockwise, while those insides are transcribed counterclockwise. The genes belonging to different functional groups are color-coded. The dark gray area of the inner circle denotes the GC content of plastome

Thunb. and *P. medicum* (Fig. 5). However, the whole plastome sequences shared low similarity among the twelve *Peucedanum* samples, identifying 7,350 variation sites in the 142,197 alignment positions (Fig. 6). According to the sequence divergences, the 15 mutation hotspot regions were selected as candidate DNA barcodes, including five protein coding genes–*ccs*A, *mat*K, *rpl*22, *rps*8, *ycf*1–which showed the Pi > 0.01200 (Fig. 7A) and 10 non-coding regions–*ccs*A-*ndh*D, *ndh*F-*rpl*32, *pet*A-*psb*J, *psb*A-*trn*K, *rpl*32-*trn*L, *rps*15-*ycf*1, *rps*2-*rpo*C2, *trn*H-*psb*A, *trn*K-*rps*16, *ycf*2-*trn*L–which showed the Pi >0.03100 (Fig. 7B).

Fig. 2 The RSCU values of all concatenated protein-coding genes for 12 *Peucedanum* plastomes. Color key: the red values represent higher RSCU values while the blue values indicate lower RSCU values

Fig. 3 Analyses of simple sequence repeats (SSRs) in twelve *Peucedanum* plastomes: **A** presence of SSRs in LSC, SSC, and IR; **B** numbers of different repeat types
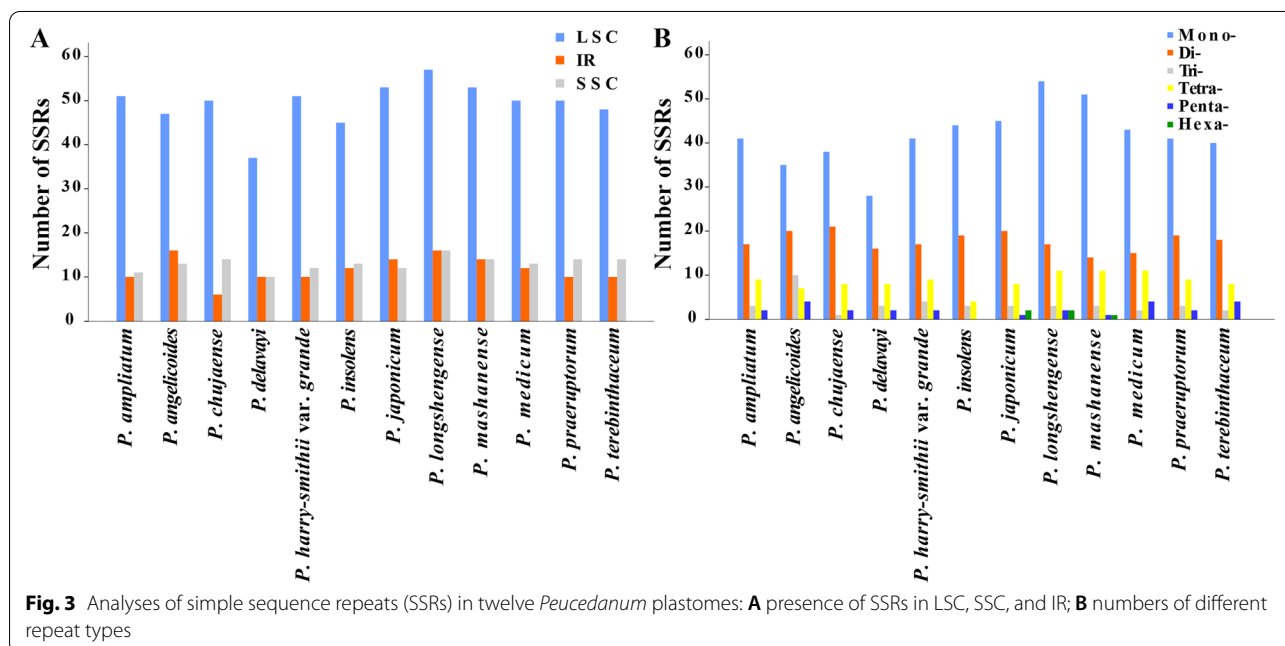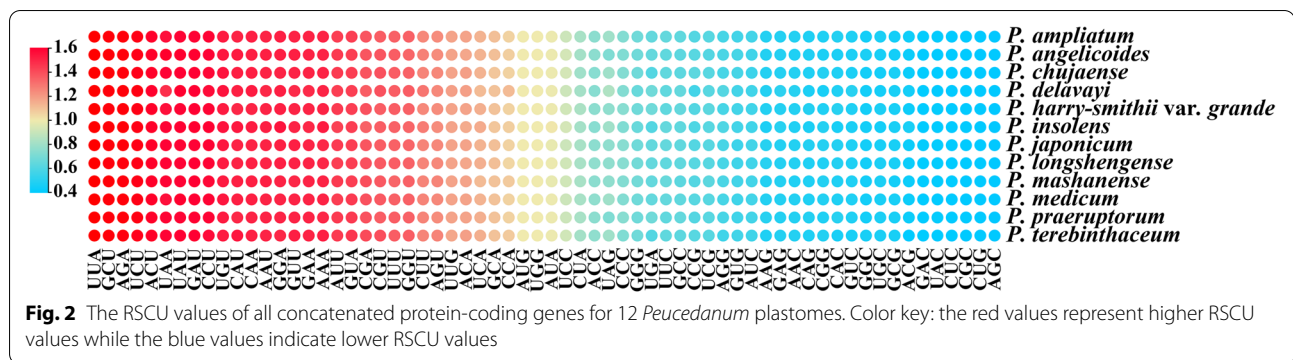
Phylogenetic analyses

The analyses of ML and BI generated the identical tree topology. The Fig. 8 illustrated the phylogeny, including two types of support values: BI posterior probabilities (PP) and ML bootstrap values (BS). Both analyses robustly supported that members of *Peucedanum* not clustered as monophyletic but fell into four clades: (1) *P. insolens* was placed in Arcuatopterus clade (PP = 1.00, BS = 100); (2) *P. delavayi* was sister to *Pterygopleurum neurophyllum* (Maxim.) Kitag., belonging to Acronema 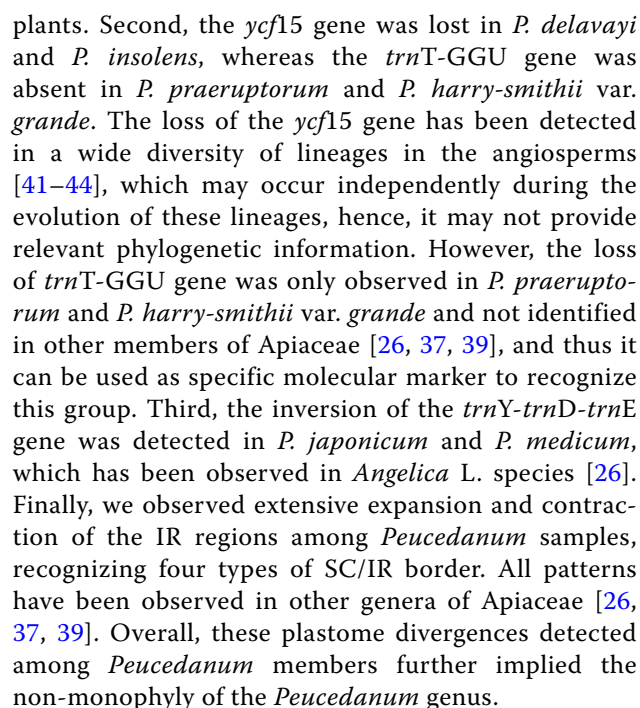clade (PP = 1.00, BS = 100); (3) *P. angelicoides* clustered with *Semenovia transiliensis* Regel & Herder constituting Tordyliinae (PP = 1.00, BS = 100); (4) the remainders were included in Selineae (PP = 1.00, BS = 100). Most of the *Peucedanum* accessions fell into the tribe Selineae, while these samples were also not clustered in a clade. Within Selineae, three major lineages

for *Peucedanum* accessions were recognized: *P. chujaense* and *P. terebinthaceum* formed a clade that was relatively distant from others (PP = 1.00, BS = 100); *P. mashanense* was clustered with *P. medicum* (PP = 1.00, BS = 100); *P. ampliatum* K.T. Fu, *P. praeruptorum*, *P. harry-smithii* var. *grande*, *P. japonicum*, and *P. longshengense* Shan et Sheh formed a clade (PP = 1.00, BS = 100), in which *P. longshengense* firstly diverged from the remainders (PP = 1.00, BS = 100), followed by *P. japonicum* (PP = 1.00, BS = 99), and the sub-clade *P. praeruptorum* + *P. harry-smithii* var. *grande* sister to *P. ampliatum* (PP = 1.00, BS = 100). In addition, the phylogenetic relationships among non-*Peucedanum* species inferred in this study were generally consistent with the previous work [37], but our results gave the higher support values for these relationships, showing PP = 1.00 and BS ≥ 96 for all nodes.

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 6 of 14

Fig. 4 Comparison of the borders of the LSC, SSC, and IR regions among twelve *Peucedanum* plastomes

Discussion

Comparison of the plastomes in *Peucedanum*

In this study, we sequenced and assembled seven plastomes of *Peucedanum* and performed a comprehensive comparative analyses of these plastomes with five other published plastomes of this genus obtained from GenBank. All *Peucedanum* plastomes showed a typically quadripartite structure, including a pair of inverted repeats regions separated by the large single copy region and small single copy region [33–36]. In addition, codon bias, RNA editing sites, and the distribution and constituent of SSRs were quite similar among twelve *Peucedanum* plastomes. These results suggested that *Peucedanum* plastome is conserved in terms of genome structure, codon bias, RNA editing sites, and SSRs. It is worth noting that this phenomenon is commonly found in other genera of flowering plants [38–40], which may be related to maintaining the stability of plastome function.

However, we also detected obvious diversity among the twelve *Peucedanum* plastomes. First, the overall sizes of plastomes varied from 142,494 bp (*P. angelicoides*) to 156,899 bp (*P. insolens*) among *Peucedanum*

plants. Second, the *ycf*15 gene was lost in *P. delavayi* and *P. insolens*, whereas the *trn*T-GGU gene was absent in *P. praeruptorum* and *P. harry-smithii* var. *grande*. The loss of the *ycf*15 gene has been detected in a wide diversity of lineages in the angiosperms [41–44], which may occur independently during the evolution of these lineages, hence, it may not provide relevant phylogenetic information. However, the loss of *trn*T-GGU gene was only observed in *P. praeruptorum* and *P. harry-smithii* var. *grande* and not identified in other members of Apiaceae [26, 37, 39], and thus it can be used as specific molecular marker to recognize this group. Third, the inversion of the *trn*Y-*trn*D-*trn*E gene was detected in *P. japonicum* and *P. medicum*, which has been observed in *Angelica* L. species [26]. Finally, we observed extensive expansion and contraction of the IR regions among *Peucedanum* samples, recognizing four types of SC/IR border. All patterns have been observed in other genera of Apiaceae [26, 37, 39]. Overall, these plastome divergences detected among *Peucedanum* members further implied the non-monophyly of the *Peucedanum* genus.

Fig. 5 Mauve alignment of twelve *Peucedanum* plastomes. Local collinear blocks within each alignment are represented by blocks of the same color connected with lines. The colored boxes are the inversion of the *trn*Y-*trn*D-*trn*E gene
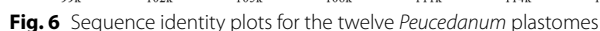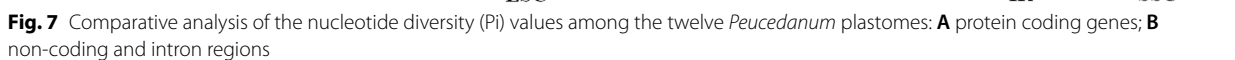
Phylogeny inference

The utilization of a small number of DNA fragments for phylogenetic analysis may cause phylogenetic errors and thus result in the incongruent topology among different DNA sequences [45–47]. Hence, using few DNA sequences to infer the phylogeny of plant species might be frequently insufficient and inappropriate, especially at low taxonomic levels [26, 47]. The plastome sequence possesses highly variable characters and thus has the tremendous potential power to reconstruct the robust phylogeny at low taxonomic levels [19–22, 31]. Therefore, we performed plastid phylogenomic analyses for *Peucedanum* genus in this study. As expected, compared to previous phylogenetic studies by using single or multiple locus DNA sequences [2, 12–16], our phylogenetic analyses based on whole plastome sequences generated a robust phylogenetic framework for *Peucedanum* members, all nodes showing PP = 1.00 and BS ≥ 96. This result justifies that the plastome sequence is powerful and effective to improve the supports and resolutions of phylogeny for *Peucedanum* genus.

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 8 of 14

Fig. 6 Sequence identity plots for the twelve *Peucedanum* plastomes

The *Peucedanum* genus was not recovered as monophyletic in our phylogenomic analyses, which was congruent with the previous studies that used ITS data and two plastid DNA regions (*rpl*16 and *rps*16 intron) [2, 12–16]. It is further supported by the great divergence of leaf epidermal morphologies [48], and fruit structures [49, 50] among *Peucedanum* members. These results justified that the *Peucedanum* genus is not a natural taxonomy unit. Therefore, the current taxonomy system of *Peucedanum* urgently needs to be improved and revised. Although the taxonomic treatment for *Peucedanum* members has not been performed in the current study due to the absence of the type species of *Peucedanum* (*P. officinale* L.), our results lay the foundations for the future taxonomic studies of *Peucedanum*.

The phylogenetic relationships among *P. japonicum*, *P. praeruptorum*, and *P. terebinthaceum* have long been controversial [14, 51, 52]. The phylogenetic analyses of Feng et al. [14] based on ITS sequences showed that *P. praeruptorum* was sister to *P. japonicum* that was relatively distant from *P. terebinthaceum*. However, the results of Ostroumova et al. [51] and Pimenov et al. [52] indicated that *P. praeruptorum* made a cluster with *P.*

terebinthaceum being sister to *P. japonicum*. Our plastid phylogenomic analyses robustly supported that *P. japonicum* was sister to the clade consisting of *P. ampliatum*, *P. praeruptorum* and *P. harry-smithii* var. *grande*, in which the subclade of *P. praeruptorum* + *P. harry-smithii* var. *grande* diverged from *P. ampliatum*; *P. terebinthaceum* and *P. chujaense* clustered into a clade that was distant from all other *Peucedanum* members. The relationships recovered in the current study are different from those of previous studies [14, 51, 52]. With high supports and resolutions, our plastid-based phylogenetic analyses provide new sights into the inter-species relationship within *Peucedanum*.
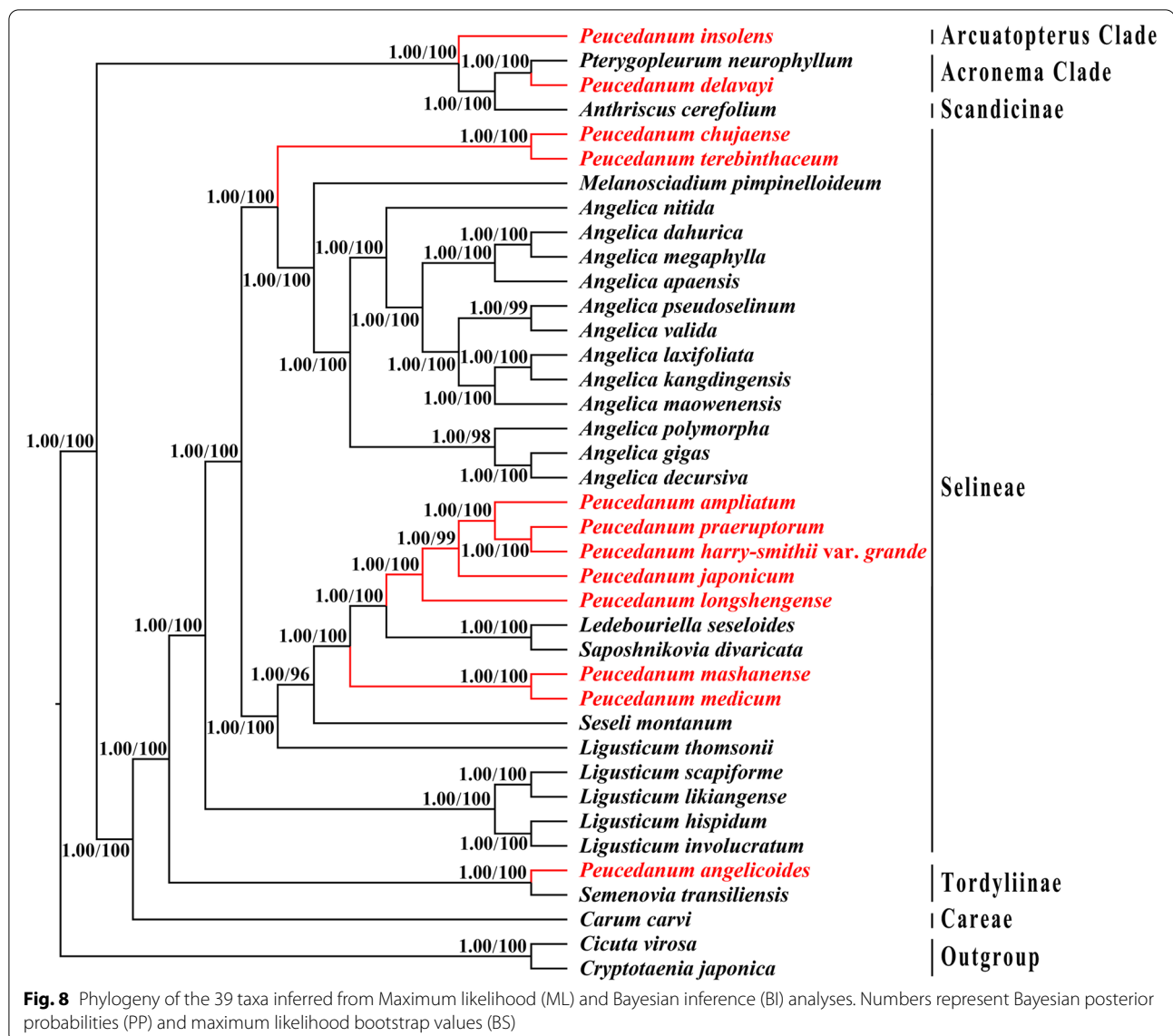
Potential DNA barcodes

The accurate species identification has always been a serious challenge faced by taxonomists. The advent of DNA barcoding technology, which uses the short DNA sequences with sufficient variations to discriminate species [53], promises to resolve this difficulty. The mitochondrial gene cytochrome oxidase 1 has been proven to be effective and reliable as a standard DNA barcode for animal species identification [54–57]. However, in plants,

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 9 of 14

Fig. 7 Comparative analysis of the nucleotide diversity (Pi) values among the twelve *Peucedanum* plastomes: **A** protein coding genes; **B** non-coding and intron regions

reliable species identification based on universal DNA barcodes, i.e., *rbc*L, *mat*K, *trn*H-*psb*A, is frequently problematic [58–62]. As expected, we found that the variation in *rbc*L gene was relatively low (Pi = 0.00925) among *Peucedanum* plants. Hence, this region may have limited power to discriminate *Peucedanum* species.

Based on sequence variations, five protein coding genes (*ccs*A, *mat*K, *rpl*22, *rps*8, *ycf*1) and ten non-coding regions (*ccs*A-*ndh*D, *ndh*F-*rpl*32, *pet*A-*psb*J, *psb*A-*trn*K, *rpl*32-*trn*L, *rps*15-*ycf*1, *rps*2-*rpo*C2, *trn*H-*psb*A, *trn*K-*rps*16, *ycf*2-*trn*L) were selected, which were potentially useful for species identification in *Peucedanum* genus. Among them, *mat*K gene and *trn*H-*psb*A region are members of universal DNA barcodes [62]; *ccs*A, *rpl*22, *ycf*1, *ccs*A-*ndh*D, *ndh*F-*rpl*32, *trn*K-*rps*16, and *ycf*2-*trn*L have been chosen as promising DNA barcodes in other plants [26, 39, 63–65]; and *pet*A-*psb*J, *rpl*32-*trn*L, and

*rps*15-*ycf*1 regions have been widely used for phylogenetic analyses [66–70]. In a future study, we will test whether or not these sequences can serve as reliable DNA barcodes for species identification within *Peucedanum* genus.

Conclusion

This study is the first attempt to comprehensively investigate the plastome features and infer phylogeny by using plastome data for *Peucedanum* genus. Comparative analyses found that plastomes of *Peucedanum* are conserved in terms of genome structure, codon bias, RNA editing sites, and SSRs, but varied in genome size, gene content and arrangement, and border of SC/IR. The plastid phylogenomic analyses prove that plastome data are efficient and powerful for improving the supports and resolutions of *Peucedanum* phylogeny and robustly support that

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 10 of 14

Fig. 8 Phylogeny of the 39 taxa inferred from Maximum likelihood (ML) and Bayesian inference (BI) analyses. Numbers represent Bayesian posterior probabilities (PP) and maximum likelihood bootstrap values (BS)

Peucedanum is not a monophyletic group. In addition, fifteen mutation hotspot regions are identified across the plastomes that can serve as potential DNA barcodes for species identification in *Peucedanum*. Overall, our study lays the foundations for the future phylogeny and taxonomy of *Peucedanum*.

Methods

Plant material, DNA extraction, plastome sequencing and assembly

The fresh young leaves of seven *Peucedanum* taxa were collected from the wild and the greenhouse in College of Life Sciences, Sichuan University, and then dried with silica gel. The formal identifications of all samples were

undertaken by Professor Xingjin He (Sichuan University). The voucher specimens were deposited at the herbarium of Sichuan University (Chengdu, China) under deposition numbers of LCK2020001- LCK2020004, LZL2020085, JQP19082303, and JQP19082505 (Table S6). Total DNA was extracted from ~20 mg silica-gel-dried leaves with the CTAB method [71]. Genomic DNA then was fragmented into 400 bp to construct the pair-end library, following the manufacturer's protocol (Illumina, San Diego, CA, USA). The libraries were sequenced on the Illumina NovaSeq platform at Personalbio (Shanghai, China). Raw data were filtered using fastP v0.15.0 (-n 10 and -q 15) to obtain high quality reads [72]. Then high-quality reads were used to assemble the whole plastome with NOVOPlasty v2.6.2 [73],

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 11 of 14

with the default parameters and *rbc*L sequence from *P. japonicum* (JF943288) as seed.

Genomic annotation and feature analyses

The assembled plastomes were annotated using web server CPGAVAS2 (http://www.herbalgenomics.org/cpgavas2) [74]. The start and stop codons and intron positions were manually corrected according to plastomes of congeneric species in Geneious v9.0.2 [75]. The maps of annotated plastomes were drawn using the online program OrganellarGenomeDRAW (OGDRAW) [76].

Five whole plastomes of *Peucedanum* (*P. chujaense*, *P. delavayi*, *P. insolens*, *P. praeruptorum*, and *P. terebinthaceum*) were downloaded from NCBI. Together with newly sequenced plastomes, we investigated the codon usage of *Peucedanum* plastomes with the CodonW v1.4.2 program [77]. Then, we predicted the potential RNA editing sites of protein coding genes for the twelve *Peucedanum* plastomes by using the online program Predictive RNA Editor for Plants suite with a cutoff value of 0.8 [78]. Moreover, simple sequence repeats (SSRs) for each plastomes were detected with MISA (http://pgrc.ipk-gatersleben.de/misa/). The thresholds of repeat units were set as 10, 5, 4, 3, 3, and 3, for mono-, di-, tri-, tetra-, penta-, and hexanucleotides, respectively.

Genomic comparison

We compared the boundaries of the LSC, SSC and IR regions among the twelve *Peucedanum* plastomes in Geneious v9.0.2 [75]. Then, the DNA rearrangements among *Peucedanum* plastomes were detected by using Mauve Alignment [79] implemented in Geneious v9.0.2 [75]. Furthermore, sequence divergence of *Peucedanum* plastomes was investigated using the mVISTA tool [80], with *P. ampliatum* set as the reference.

Identification of divergence hotspots

In order to identify mutation hotspot regions, the protein coding genes, non-coding regions and intron regions of the twelve *Peucedanum* plastomes were extracted in Geneious v9.0.2 [75] and aligned with MAFFT v7.221 [81]. Then, alignments with more than 200 bp in length were used to evaluate nucleotide diversity (Pi) using DnaSP v5.0 [82]. The thresholds of Pi for protein coding gene and non-coding region were set as 0.01200 and 0.03100, respectively.

Phylogenetic analyses

To infer the phylogenetic relationships among *Peucedanum* species, we reconstructed phylogenetic trees using 39 plastomes (Table S6, Table S7). *Cicuta virosa* L. and *Cryptotaenia japonica* Hassk. were chosen as outgroup to root the phylogenetic tree, according to the results of Wen et al. [37]. Sequence alignment was performed with the software MAFFT v7.221 [81], and adjusted and corrected manually when necessary. The unambiguous matrix was subjected to Maximum-Likelihood analyses (ML) and Bayesian Inference (BI). The ML phylogenetic tree was reconstructed in the program RAxML v8.2.8 [83] with 1000 replicates and GTRGAMMA model as the RAxML manual suggested. The BI analysis was performed by using MrBayes v3.2.7 [84] with the best-fit substitution model (TVM+I+G) determined by Modeltest v3.7 [85]. Two independent Markov chains were run for 1,000,000 generations, sampling every 100 generations. The first 25% of trees were discarded as burn-in and the remainder were used to generate the consensus tree. Results of phylogenetic analyses were visualized and edited in FigTree v1.4.2 [86].

Abbreviations
BI: Bayesian inference; bp: Base pair; BS: Bootstrap value; CTAB: Cetyl trimethylammonium bromide; IR: Inverted repeat; ITS: Internal transcribed spacer; LSC: Large single copy; ML: Maximum Likelihood; Pi: Nucleotide diversity; PP: Posterior probability; rRNA: Ribosomal RNA; RSCU: Relative synonymous codon usage; SSC: Small single copy; SSR: Simple sequence repeat; tRNA: Transfer RNA.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12870-022-03488-x.

Additional file 1: Fig. S1. Analyses of RNA editing sites in twelve *Peucedanum* plastomes: (A) numbers of RNA editing sites distributed in different codon positions; (B) numbers of RNA editing sites presented in genes.

Additional file 2: Table S1. Summary of Illumina sequencing of *Peucedanum*.

Additional file 3: Table S2. List of unique genes identified in plastomes of *Peucedanum*.

Additional file 4: Table S3. Codon usage and relative synonymous codon usage (RSCU) values of protein-coding genes of the twelve *Peucedanum* plastomes.

Additional file 5: Table S4. RNA editing sites detected in the twelve plastome of *Peucedanum*.

Additional file 6: Table S5. Numbers of SSR motifs identified in the twelve *Peucedanum* plastomes.

Additional file 7: Table S6. Taxa newly sequenced in the present study with source, voucher and GenBank accession numbers.

Additional file 8: Table S7. Plastomes included in phylogenetic analyses with GenBank accession.

Authors' contributions
S-DZ and X-JH designed the research. C-KL, J-QL, and Q-PJ collected and analyzed the data; C-KL, S-DZ and X-JH prepared the manuscript. All authors read and approved the final manuscript.

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 12 of 14

Declarations

Ethics approval and consent to participate
Collection of all samples completely complies with national and local legislation permission. Plant samples used in the study were not included in the list of national key protected plants and not collected from national park or natural reserve. According to national and local legislation, no specific permission was required for collecting these plants.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References
1. Sheh ML. *Peucedanum*. In: Shan RH, Sheh ML, editors. Flora Reipublicae Popularis Sinica. Beijing: Science Press; 1992. p. 123–75.
2. Spalik K, Reduron JP, Downie SR. The phylogenetic position of *Peucedanum sensu lato* and allied genera and their placement in tribe Selineae (Apiaceae, subfamily Apioideae). Pl Syst Evol. 2004;243(3):189–210.
3. Sheh ML, Watson MF. *Peucedanum* Linnaeus. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China. Beijing: Science Press & St. Louis: Missouri Botanic Garden Press; 2005. p. 182–92.
4. Drude CGO. Umbelliferae. In: Engler A, Prantl K, editors. Die naturlichen Pflanzenfamilien. Leipzig: Wilhelm Engelmann; 1897–1898. p. 63–250.
5. Pimenov MG, Leonov MV. The genera of Umbelliferae. Kew: Royal Botanical Gardens; 1993.
6. Wu ZY, Lu AM, Tang YC, Chen ZD, Li DZ. The families and genera of Angiosperms in China: a comprehensive analysis. Beijing: Science Press; 2003.
7. Zhang X. Studies on geographic distribution of *Peucedanum* L. in southwest of China. Chin Agric Sci Bull. 2011;27(12):177–80.
8. Downie SR, Spalik K, Katz-Downie DS, Reduron JP. Major clades within Apiaceae subfamily Apioideae as inferred by phylogenetic analysis of nrDNA ITS sequences. Plant Divers Evol. 2010;128(1):111–36.
9. Shneyer VS, Kutyavina NG, Pimenov MG. Systematic relationships within and between *Peucedanum* and *Angelica* (Umbelliferae–Peucedaneae) inferred from immunological studies of seed proteins. Plant Syst Evol. 2003;236(3):175–94.
10. Reduron JP, Charpin A, Pimenov MG. Contribution a` la nomenclature ge´ne´rique des Apiaceae (Ombellife` res). J Bot Soc Bot Fr. 1997;1:91–104.
11. Winter PJD, Magee AR, Phephu N, Tilney PM, Downie SR, van Wyk BE. A new generic classification for African peucedanoid species (Apiaceae). Taxon. 2008;57(2):347–64.
12. Downie SR, Watson MF, Spalik K, Katz-Downie DS. Molecular systematics of Old World Apioideae (Apiaceae): relationships among some members of tribe Peucedaneae sensu lato, the placement of several island-endemic species, and resolution within the apioid superclade. Can J Bot. 2000;78(4):506–28.
13. Valiejo-Roman CM, Terentieva EI, Samigullin TH, Pimenov MG, Ghahremani-Nejad F, Mozaffarian V. Molecular data (nrITS-sequencing) reveal relationships among Iranian endemic taxa of the Umbelliferae. Feddes Repert. 2006;117(5-6):367–88.
14. Feng T, Downie SR, Yu Y, Zhang XM, Chen WW, He XJ, et al. Molecular systematics of *Angelica* and allied genera (Apiaceae) from the Hengduan Mountains of China based on nrDNA ITS sequences: phylogenetic affinities and biogeographic implications. J Plant Res. 2009;122(4):403–14.
15. Zhou J, Gong X, Downie SR, Peng H. Towards a more robust molecular phylogeny of Chinese Apiaceae subfamily Apioideae: additional evidence from nrDNA ITS and cpDNA intron (*rpl*16 and *rps*16) sequences. Mol Phylogenet Evol. 2009;53(1):56–68.
16. Zhou J, Gao YZ, Wei J, Liu ZW, Downie SR. Molecular phylogenetics of *Ligusticum* (Apiaceae) based on nrDNA ITS sequences: rampant polyphyly, placement of the Chinese endemic species, and a much-reduced circumscription of the genus. Int J Plant Sci. 2020;181(3):306–23.
17. Kim K, Kim CS, Oh SH, Park CW. A new species of *Peucedanum* (Apiaceae) from Korea. Phytotaxa. 2019;393(1):75–83.
18. Sarkhail P. Traditional uses, phytochemistry and pharmacological properties of the genus *Peucedanum*: a review. J Ethnopharmacol. 2014;156:235–70.
19. Wicke S, Schneeweiss GM, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 2011;76(3):273–97.
20. Shaw J, Shafer HL, Leonard OR, Kovach MJ, Schorr M, Morris AB. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms, the tortoise and the hare IV. Am J Bot. 2014;101(11):1987–2004.
21. Tonti-Filippini J, Nevill PG, Dixon K, Small I. What can we do with 1000 plastid genomes? Plant J. 2017;90:808–e818.
22. Schneider JV, Paule J, Jungcurt T, Cardoso D, Amorim AM, Berberich T, et al. Resolving recalcitrant clades in the pantropical Ochnaceae: insights from comparative phylogenomics of plastome and nuclear genomic data derived from targeted sequencing. Front Plant Sci. 2021;12:638650.
23. Ravi V, Khurana JP, Tyagi AK, Khurana P. An update on chloroplast genomes. Plant Syst Evol. 2008;271(1):101–22.
24. Guisinger MM, Kuehl JV, Boore JL, Jansen RK. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. Mol Biol Evol. 2011;28(1):583–600.
25. Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, et al. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci U S A. 2007;104(49):19369–74.
26. Wang M, Wang X, Sun J, Wang Y, Ge Y, Dong W, et al. Phylogenomic and evolutionary dynamics of inverted repeats across *Angelica* plastomes. BMC Plant Biol. 2021;21:26.
27. Nock CJ, Waters DLE, Edwards MA, Bowen SG, Rice N, Cordeiro GM, et al. Chloroplast genome sequences from total DNA for plant identification. Plant Biotechnol J. 2011;9(3):328–33.
28. Ruhsam M, Rai HS, Mathews S, Ross TG, Graham SW, Raubeson LA, et al. Does complete plastid genome sequencing improve species discrimination and phylogenetic resolution in *Araucaria*? Mol Ecol Resour. 2015;15(5):1067–78.
29. Ji Y, Yang L, Chase MW, Liu C, Yang Z, Yang J, et al. Plastome phylogenomics, biogeography, and clade diversification of *Paris* (Melanthiaceae). BMC Plant Biol. 2019;19(1):543.
30. Li HT, Yi TS, Gao LM, Ma PF, Zhang T, Yang JB, et al. Origin of angiosperms and the puzzle of the Jurassic gap. Nat Plants. 2019;5(5):461–70.
31. Liu C, Yang J, Jin L, Wang S, Yang Z, Ji Y. Plastome phylogenomics of the East Asian endemic genus *Dobinea*. Plant Divers. 2021;43(1):35–42.
32. Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol. 2009;7(1):84.
33. Lee HO, Joh HJ, Kim K, Lee SC, Kim NH, Park JY, et al. Dynamic chloroplast genome rearrangement and DNA barcoding for three Apiaceae species known as the medicinal herb "Bang-Poong". Int J Mol Sci. 2019;20(9):2196.
34. Li Y, Geng M, Xu Z, Wang Q, Li L, Xu M, et al. The complete plastome of *Peucedanum praeruptorum* (Apiaceae). Mitochondrial DNA B Resour. 2019;4(2):3612–3.
35. Gou W, Jia SB, Price M, Guo XL, Zhou SD, He XJ. Complete plastid genome sequencing of eight species from *Hansenia*, *Haplosphaera* and *Sinodielsia* (Apiaceae): comparative analyses and phylogenetic implications. Plants. 2020;9(11):1523.

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 13 of 14

36. Kim K, Lim CE, Lee BY. The complete chloroplast genome sequence of *Peucedanum chujaense* (Apiaceae), an endemic species to Korea. Mitochondrial DNA B Resour. 2020;5(2):1937–8.

37. Wen J, Xie DF, Price M, Ren T, Deng Y, Gui L, et al. Backbone phylogeny and evolution of Apioideae (Apiaceae): new insights from phylogenomic analyses of plastome data. Mol Phylogenet Evol. 2021;161:107183.

38. Mwanzia VM, He DX, Gichira AW, Li Y, Ngarega BK, Karichu MJ, et al. The complete plastome sequences of five *Aponogeton* species (Aponogetonaceae): insights into the structural organization and mutational hotspots. Plant Divers. 2020;42(5):334–42.

39. Ren T, Li ZX, Xie DF, Gui LJ, Peng C, Wen J, et al. Plastomes of eight *Ligusticum* species: characterization, genome evolution, and phylogenetic relationships. BMC Plant Biol. 2020;20(1):519.

40. Mehmood F, Rahim A, Heidari P, Ahmed I, Peter P. Comparative plastome analysis of *Blumea*, with implications for genome evolution and phylogeny of Asteroideae. Eco Evolution. 2021;11(12):7810–26.

41. Nguyen PAT, Kim JS, Kim JH. The complete chloroplast genome of colchicine plants (*Colchicum autumnale* L. and *Gloriosa superba* L.) and its application for identifying the genus. Planta. 2015;242(1):223–37.

42. Shin DH, Lee JH, Kang SH, Ahn BO, Kim CK. The complete chloroplast genome of the hare's ear root, *Bupleurum falcatum*: its molecular features. Genes. 2016;7(5):20.

43. Sun L, Jiang Z, Wan X, Zou X, Yao X, Wang Y, et al. The complete chloroplast genome of *Magnolia polytepala*: comparative analyses offer implication for genetics and phylogeny of Yulania. Gene. 2020;736:144410.

44. Guo X, Liu C, Wang H, Zhang G, Yan H, Jin L, et al. The complete plastomes of two flowering epiparasites (*Phacellaria glomerata* and *P. compressa*): gene content, organization, and plastome degradation. Genomics. 2021;113(2):447–55.

45. Rokas A, Carroll SB. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol Biol Evol. 2005;22:1337–44.

46. Philippe H, Brinkmann H, Lavrov DV, Littlewood DTJ, Manuel M, Wörheide G, et al. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 2011;9:e1000602.

47. Huang Y, Li X, Yang Z, Yang C, Yang J, Ji Y. Analysis of complete chloroplast genome sequences improves phylogenetic resolution in *Paris* (Melanthiaceae). Front Plant Sci. 2016;7:1797.

48. Zhou J, Wang WC, Gong X, Liu ZW. Leaf epidermal morphology in *Peucedanum* L. (Umbelliferae) from China. Acta Bot Gallica. 2014;161(1):21–31.

49. Wang MY, Liu M, Cheng XY, Wang YT. The phylogenetic significance of fruit structures in Chinese *Peucedanum* of Apiaceae. Acta Pratac Sin. 2015;24(6):168–76.

50. Ostroumova TA. Fruit micromorphology in the Umbelliferae of the Russian Far East. Bot Pac. 2018;7(1):41–9.

51. Ostroumova TA, Pimenov MG, Degtjareva GV, Samigullin TH. *Taeniopetalum* Vis. (Apiaceae), a neglected segregate of Peucedanum L., supported as a remarkable genus by morphological and molecular data. Skvortsovia. 2016;3:20–44.

52. Pimenov MG, Ostroumova TA, Degtjareva GV, Samigullin TH. *Sillaphyton*, a new genus of the Umbelliferae, endemic to the Korean Peninsula. Bot Pac. 2016;5(2):31–41.

53. Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. Plant DNA barcoding: from gene to genome. Biol Rev. 2015;90(1):157–66.

54. Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN. DNA barcodes distinguish species of tropical Lepidoptera. Proc Natl Acad Sci U S A. 2006;103(4):968–71.

55. Costa FO, DeWaard JR, Boutillier J, Ratnasingham S, Dooh RT, Hajibabaei M, et al. Biological identifications through DNA barcodes: the case of the Crustacea. Can J Fish Aquat Sci. 2007;64(2):272–95.

56. Smith MA, Poyarkov NA, Hebert PDN. DNA barcoding: CO1 DNA barcoding amphibians: take the chance, meet the challenge. Mol Ecol Resour. 2008;8(2):235–46.

57. Kim DW, Yoo WG, Park HC, Yoo HS, Kang DW, Jin SD, et al. DNA barcoding of fish, insects, and shellfish in Korea. Genomics Inform. 2012;10(3):206–11.

58. Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, et al. A DNA barcode for land plants. Proc Natl Acad Sci U S A. 2009;106(31):12794–7.

59. Hollingsworth PM. Refining the DNA barcode for land plants. Proc Natl Acad Sci U S A. 2011;108(49):19451–2.

60. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. PLoS One. 2011;6(5):e19254.

61. Coissac E, Hollingsworth PM, Lavergne S, Taberlet P. From barcodes to genomes: extending the concept of DNA barcoding. Mol Ecol. 2016;25:1423–8.

62. Hollingsworth PM, Li DZ, van der Bank M, Twyford AD. Telling plant species apart with DNA: from barcodes to genomes. Philos Trans R Soc B. 2016;371(1702):20150338.

63. Liu C, Yang Z, Yang L, Yang J, Ji Y. The complete plastome of *Panax stipuleanatus*: Comparative and phylogenetic analyses of the genus *Panax* (Araliaceae). Plant Divers. 2018;40(6):265–76.

64. Xie DF, Yu HX, Price M, Xie C, Deng YQ, Chen JP, et al. Phylogeny of Chinese *Allium* species in section *Daghestanica* and adaptive evolution of *Allium* (Amaryllidaceae, Allioideae) species revealed by the chloroplast complete genome. Front Plant Sci. 2019;10:460.

65. Dong W, Liu Y, Xu C, Gao Y, Yuan Q, Suo Z, et al. Chloroplast phylogenomic insights into the evolution of *Distylium* (Hamamelidaceae). BMC Genomics. 2021;22(1):293.

66. Ngamriabsakul C, Techaprasan J. The phylogeny of Thai Boesenbergia (Zingiberaceae) based on *pet*A-*psb*J spacer (chloroplast DNA). J Sci Technol. 2006;28:49–57.

67. Jaramillo MA, Callejas R, Davidson C, Smith JF, Stevens AC, Tepe EJ. A phylogeny of the tropical genus *Piper* using ITS and the chloroplast intron *psb*J-*pet*A. Syst Bot. 2008;33(4):647–60.

68. Falchi A, Paolini J, Desjobert JM, Melis A, Costa J, Varesi L. Phylogeography of *Cistus creticus* L. on Corsica and Sardinia islands inferred by the *trn*L-F and *rpl*32-*trn*L sequences of cpDNA. Mol Phylogenet Evol. 2009;52(2):538–43.

69. Yang JB, Yang HQ, Li DZ, Wong KM, Yang YM. Phylogeny of *Bambusa* and its allies (Poaceae: Bambusoideae) inferred from nuclear GBSSI gene and plastid *psb*A-*trn*H, *rpl*32-*trn*L and *rps*16 intron DNA sequences. Taxon. 2010;59(4):1102–10.

70. Chen YP, Drew BT, Li B, Soltis DE, Soltis PS, Xiang CL. Resolving the phylogenetic position of *Ombrocharis* (Lamiaceae), with reference to the molecular phylogeny of tribe Elsholtzieae. Taxon. 2016;65(1):123–36.

71. Doyle JJ, Doyle JL. Arapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.

72. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90.

73. Dierckxsens N, Mardulyn P, Smits G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 2017;45(4):e18.

74. Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, et al. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 2019;47(W1):W65–73.

75. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9.

76. Lohse M, Drechsel O, Bock R. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet. 2007;52:267–74.

77. Peden JF. Analysis of codon usage. PhD thesis. Nottingham: University of Nottingham; 1999.

78. Mower JP. The PREP suite: predictive RNA editors for plant mitochondrial genes, chloroplast genes, and user-defined alignments. Nucleic Acids Res. 2009;37:W253–9.

79. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004;14(7):1394–403.

80. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 2004;32(suppl_2):W273–9.

81. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.

82. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25:1451–2.

Liu *et al. BMC Plant Biology*   (2022) 22:101

Page 14 of 14

83. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–3.

84. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–42.

85. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics. 1998;14(9):817–8.

86. Rambaut A, Drummond A. FigTree, version 1.4.2. 2015. http://tree.bio.ed.ac.uk/software/figtree/. Accessed 4 Oct 2021.

Publisher's Note