Pak. J. Bot., 57(5): DOI: http://dx.doi.org/10.30848/PJB2025-5(1)

COMPLETE CHLOROPLAST GENOME OF *PREMNA OBTUSIFOLIA*, A SEMI MANGROVE PLANT

SHI-QUAN WANG

College of Biology and Food Engineering, Fuyang Normal University, Fuyang 236037, China
Corresponding email: wsqmah@163.com

Abstract

Premna obtusifolia is one of distributed semi-mangrove species in the world. In this research, whole chloroplast (cp) genome of *P. obtusifolia* was assembled for the first time. Meanwhile, this research expanded the chloroplast genome resource bank of semi-mangrove species. Comprehensive analysis about cp genome of *P. obtusifolia* showed prospective cpDNA markers, which could be used to study species identification, species adaptation mechanisms and phylogenetic status. The total cp genome was 153,415 bp in length, displaying paradigmatic tetrad structure, which included a large single copy (LSC) region of 84,649bp and a small single copy (SSC) region of 17,410bp, and they were segregated by two inverted repeats (IRs) regions of 25,678bp. Overall GC content was 38.02% of cp genome, and GC content in IRs, LSC, and SSC regions was 43.21%, 36.03%, and 32.38%, respectively. The genome was predicted through 114 unique genes, embodying 80 protein-coding genes, 30 tRNA genes, and 4 rRNA genes, and there were 19 duplicates in IR regions. 87 SSRs and 17 long repeats were recognized totally. Phylogenetic analysis suggested that *P. obtusifolia* was clustered with *P. vietnamensis* within genus *Premna*. This research shows that the cp genome of *P. obtusifolia* is conservative in structure with slow evolutionary rate and is more structurally differentiated compared with other species in *Premna*. These results indicate that semi-mangrove species respond to the disturbance of the external environment well in the evolutionary history, and their genome characteristics have been preserved.

Key words: *Premna obtusifolia*, chloroplast, phylogenetic analysis, comparative genome, evolution

Introduction

Chloroplast (cp) is an organelle with multi-function where the photosynthesis happened in green plants, and plays a key role in fixation carbon and light energy conversion into chemical energy (Daniell, 2016; Mo *et al*., 2020). Chloroplast is a circular DNA molecule of double-strand that is maternally inherited, semi-autonomous organelle, and has relatively steady structure (Drouin *et al*., 2008). Chloroplast genome in angiosperms usually consists of 2 inverted repeats (IR region, 15-30kb), one short single-copy sequence (SSC) and one long single-copy sequence (LSC). Generally, cp genome size of land plants is 120-160 kb, encoding 110-130 unique genes (Huang *et al*., 2014). Due to small size of the genome, maternal inheritance, slow evolution rate and little recombination in most angiosperms, it is a good material for studying species evolution and classification (Ando *et al*., 2018).

Mangrove and semi-angrove plants grow in tropical and subtropical coastal intertidal zones and are the last ecological barrier between land and sea (Guo *et al*., 2017). They are distributed along the coast of south regions in China. *Premna obtusifolia* belongs to Lamiaceae and distributed in scattered patterns. It is one of the components of coastal shrubs, with a small number of individuals in the wild, and is classified as a vulnerable species (Wang *et al*., 2016).

In this investigation, leaves of *P. obtusifolia* were harvested, and total DNA was isolated. By high-throughput sequencing and assembly, a complete cp genome sequence of typical tetrad structure was obtained. The comparative analysis explored the structure variation and evolution of *P. obtusifolia*. This research enriched chloroplast information of mangrove species and provided a research basis for its resource protection.

Material and Methods

Material collection and chloroplast genome sequencing: The leaves of *P. obtusifolia* used in this research were gathered in Wenchang, Hainan, China (19.625685°N, 110.814424°E). About 10 g fresh leaves was harvested for cpDNA extraction by improved separation method (Daniell, 2016). DNA quality was evaluated through Qubit 3.0 and gel electrophoresis on 1% agarose. 1 μg of high quality DNA (OD260/OD280 = 1.8~2.0, > 6 ug) was employed to assemble short-insert library (insert size 400 bp) based on manufacturer's instruction (Illumina), and sequenced through Illumina Hiseq 4000 (Drouin *et al*., 2008).The formal identification of the species used in this research was conducted by Shi-Quan Wang.

Assembly and annotation of chloroplast genome: The initial assembly of cp genome was performed by SPAdes software (v 2.04) (Bankevich *et al*., 2012). Later, contigs were blasted to reference cp genome in *Premna*, and aligned contigs of high similarity (≥ percent 80) were sorted according to reference genome. Gaps and false bases of cp genome sketch in *P. obtusifolia* were corrected by Pilon software (v1.12) (Walker *et al*., 2014). Lastly, a circular topology cp genome of *P. obtusifolia* was obtained.

The online DOGMA means (Wyman *et al*., 2004) were employed to annotate cp genes of *P. obtusifolia*, including predicting protein-coding genes, transfer RNA (tRNA) genes and ribosome RNA (rRNA) genes. In addition, tRNA genes were further identified through tRNAscan-SE (Lowe & Eddy, 1997). Wide cp genome Blast (Altschul *et al*., 1990) research was conducted on KEGG (Kanehisa *et al*., 2006; Kanehisa *et al*., 2004) and NR (Non-Redundant Protein Database). A circular plot of cp genome in *P. obtusifolia* was created by OrganellarGenomeDRAW program (Lohse *et al*., 2007). Complete cp genome sequence of *P. obtusifolia* has been stored in GenBank (accession number PP103654).

Repetitive sequences and codon usage analysis: Simple sequence repeat (SSR) was recognized through MISA software (Beier *et al*., 2017) with arguments as followed: at least 8 repeat units of mono-nucleotides, at least 5 repeat units of di-nucleotides, at least 4 repeat units of tri-nucleotides, and at least 3 repeat units of tetra-, penta-, and hexa-nucleotides (Yi *et al*., 2013). Long repetitive sequences containing forward, reverse, complementary and tetrad repeats were analyzed through online REPuter software (Kurtz & Schleiermacher, 1999; Kang *et al*., 2019), with minimum 30 bp repeat size and three hamming distance.

Relative Synonymous Codon Usage (RSCU) was a key analysis for conducting synonymous codon usage analysis (Sablok *et al*., 2011). In this investigation, RSCU result of protein-coding genes in *P. obtusifolia* was analyzed by CodonW v1.4.2 program (Xu *et al*., 2011).

Comparison of chloroplast genome: Cp genome of *P. obtusifolia* was compared with that of *P. vietnamensis* (MT473774), *P. szemaoensis* (MT473775), *P. microphylla* (NC_026291) and *P. puberula* (NC_061379) through mVISTA tool by Shuffle-LAGAN mode (Frazer *et al*., 2004). Some species have also been employed to compare boundaries of LSC/IRB/SSC/IRA region with *P. obtusifolia* as control. The 80 protein-coding genes from the species were aligned through MAFFT v7.309 (Katoh *et al*., 2002). Analysis of maximal likelihood (ML) bootstrap with 1000 replicates was employed using MEGA7 (Kumar *et al*., 2016).

Results

Chloroplast genome feature of *P. obtusifolia*: Chloroplast genome of *P. obtusifolia* is a round double-chain DNA molecule, whose total length is 153,415 bp. Just as other angiosperms, the round cp genome of *P. obtusifolia* exhibits a paradigmatic tetrad structure with one LSC (84,649 bp), one SSC (17,410 bp) and two inverted repeat IR regions (IRa and IRb, each 25,678 bp) (Fig. 1 and Table 1). In general, GC content (38.02%) of cp genome in *P. obtusifolia* is relatively low. After evaluating GC contents of LSC, SSC and IR regions, GC content (43.21%) of IR regions was greater than that in LSC (36.03%) and SSC (32.38%), which seemed to be a common phenomenon (Zhang *et al*., 2019; Chen *et al*., 2018; Nazareno *et al*., 2015). This may be owing to relative high GC-content of rRNA and tRNA genes (Gao *et al*., 2009), and their occupation of a larger region than protein-coding genes in IR regions (Curci *et al*., 2015).

Cp genome of *P. obtusifolia* was annotated to encode 133 genes, with 114 unique genes (including 80 protein-coding genes, 30 tRNAs and 4 rRNAs) and 19 duplicated genes in IR regions (Table 1). In 19 duplicated genes, 6 were protein-coding genes, 9 were tRNAs and 4 were rRNAs (Table 2). As is common in other terrestrial plants (Redwan *et al*., 2015; Hu *et al*., 2017), 19 genes containing intron existed in *P. obtusifolia* cp genome, cluding 6 tRNA and 13 protein-coding genes, with 15 genes being composed of single intron and 4 genes (*rps12*, *ycf3*, *accD* and *clpP*) including two introns (Table 2).

The intron in trnK-UUU, including the *matK* gene (Ohsako & Ohmi, 2001), was the longest gene in all genes, amount to 2,578 bp. As former investigation (Yang *et al*., 2013; Wang *et al*., 2017), *rps12* is a reverse splicing gene, with 1 exon locating in LSC region (5′end) and the other two exons (segregated by an intron) locating in two IR regions.

Table 1. Summary of the *P. obtusifolia* chloroplast genome features.

Genome features	*P. obtusifolia*
Genome size (bp)/GC content (%)	153,415 / 38.02
LSC size (bp)/GC content (%)	84,649 / 36.03
SSC size (bp)/GC content (%)	17,410 / 32.38
IR size (bp)/GC content (%)	25,678 / 43.21
Total gene number	133
Unique gene number	114
Protein-coding gene	80
tRNAs	30
rRNAs	4
Genes duplicated in IR	19

Analyses of SSRs and long repeats: SSRs were extensively distributed in cp genome and has been widely employed in research of population genetics and molecular phylogenetics (Pervaiz *et al*., 2015). In the investigation, 87 SSRs were recognized in *P. obtusifolia* chloroplast genomes by MISA (Table 3). There were 26 mononucleotides (29.89%), 30 dinucleotides (34.48%), 13 trinucleotides (44.94%), 16 tetranucleotides (18.39%), and 2 pentanucleotide (2.3%). Additionally, A/T repeat sequences were the most common of mono-nucleotides (100%), and AA/TT repeats were majority of di-nucleotide repeat sequences (86.67%). Normally, the percentage would be reduced by the size of SSR increase, however, the quantity of tetra-repeat is more than tri-repeat. Long repeats of cp genomes were regarded as uncommon for most terrestrial plants (Mo *et al*., 2020; Huang *et al*., 2014). A total of 17 long repetitive sequences were recognized in *P. obtusifolia* cp genomes, including 7 forward repetitive sequences and 10 palindromic repetitive sequences (Table 4). The repeats with the size more than 62 bp were found to be the most common in *P. obtusifolia*. Furthermore, it contains 12 repeats with the size over 100000 bp. It means that the quantity of repeats was not only decided by the repeat size.

Analysis of codon preference: An amino acid was encoded by at least a codon (synonymous codon) in an organism (Mo *et al*., 2020; Young & Purton, 2016), and this phenomenon is considered to be codon's degeneracy. The codon's degeneracy is important for plants, because it can decrease impact of harmful mutation in genetics. The phenomenon of synonymous codons generally showed preference in the process of plants' evolution (Liu & Xue, 2005). The relative synonymous codon usage (RSCU) is an efficacious parameter to discover codon's preference (Sharp & Li, 1987). In *P. obtusifolia* cp genome, there existed 20,903 codons, 64 codons encoding 20 amino acids (Table 4). Totally, there are 30 codons, whose RSCU values are above 1, especially, 24 codons showed little preference (RSCU values = 1~1.6), 3 codons were median preference (RSCU values = 1.6~1.7), and RSCU values of 3 codons ranged from 1.7 to 2.0 (Table 5) (Yu *et al*., 2018; Zuo *et al*., 2017). Codons exhibiting preference could be a cause for relative conservation of cp genomes (Wu *et al*., 2018).

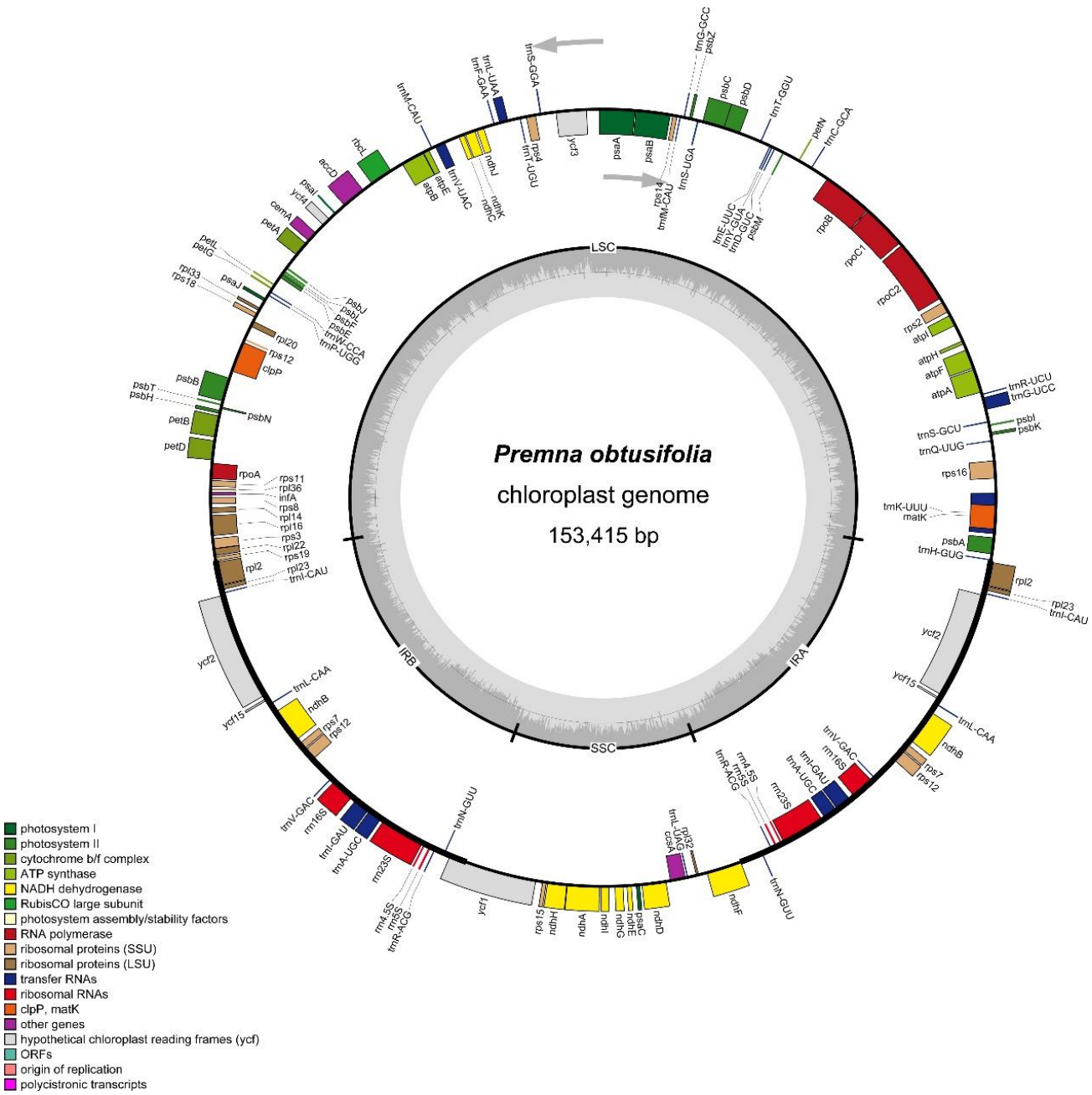

Fig. 1. Gene map of *P. obtusifolia* chloroplast genome. Genes drawn inside and outside of the circle are transcribed in the clockwise and counterclockwise directions, respectively. Genes belonging to different functional groups are color coded. The darker and lighter gray in the inner circle different functional groups are color coded. The darker and lighter gray in the inner circle corresponds corresponds to GC and AT content, respectively. LSC, large single copy region; SSC, small single copy region; IR, inverted repeat.

Analysis of comparative chloroplast genome: Wide chloroplast genome sequence of *P. obtusifolia* (PP103654), as a reference, was compared with those of *P. vietnamensis* (MT473774), *P. szemaoensis* (MT473775), *P. microphylla* (NC_026291), and *P. puberula* (NC_061379) using mVISTA program (Fig. 2). Comparison results indicated that the category, quantity, and arrangement order of genes encoded by the cp genome sequence of species in *Premna* are highly consistent; The variation between sequences primarily occurs in non-coding intergenic region. The variation between species is mainly reflected in gene regions, for example, *rps16-trnQ-UUG, atpH-atpI, petN-psbM, rpl20-clpP, rpl32-trnL-UAG, ycf3-trnS-GGA, clpP, petB, ycf1, ycf2*, etc (Fig. 2).

Contraction and expansion of IR regions: The contraction and expansion of IR regions could measure discrepancy among cp genomes (Chen *et al*., 2018; He *et al*., 2020). Comparison of IR/SC connections between *P. obtusifolia* (PP103654), *P. vietnamensis* (MT473774), *P. szemaoensis* (MT473775), *P. microphylla* (NC_026291) and *P. puberula* (NC_061379) were presented in Fig. 3. Compared with other species, *ndhF* gene of *P. microphylla* is situated in SSC region, indicating that SSC region of *P. microphylla* expanded. Compared to other species, the *ycf1* gene of *P. puberula* is shorter, with *ycf1* fragment located in IRa region of only 270 bp. Near JLA boundary, there is a gene, *rps19* in *P. vietnamensis* and *P. szemaoensis*, while there is another gene *rpl2* in other species (Fig. 3).

Table 2. List of genes annotated in the chloroplast genomes of *P. obtusifolia*.

Category	Gene group	Gene name	Number
Photosynthesis	Subunits of photosystem I	*psaA,psaB,psaC,psaI,psaJ*	5
	Subunits of photosystem II	*psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ*	15
	Subunits of NADH dehydrogenase	*ndhA*, ndhB*(2), ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK*	11(1)
	Subunits of cytochrome b/f complex	*petA, petB*, petD*, petG, petL, petN*	6
	Subunits of ATP synthase	*atpA, atpB, atpE, atpF*, atpH, atpI*	6
	Large subunit of rubisco	*rbcL*	1
Self-replication	Proteins of large ribosomal subunit	*rpl14, rpl16*, rpl2*, rpl20, rpl22, rpl23, rpl32, rpl33, rpl36*	9
	Proteins of small ribosomal subunit	*rps11, rps12**(2), rps14, rps15, rps16*, rps18, rps19, rps2, rps3, rps4, rps7(2), rps8*	12(2)
	Subunits of RNA polymerase	*rpoA, rpoB, rpoC1*, rpoC2*	4
	Ribosomal RNAs	*rrn16S(2), rrn23S(2), rrn4. 5S(2), rrn5S(2)*	4(4)
	Transfer RNAs	*trnA-UGC*(2), trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC*, trnH-GUG(2), trnI-CAU(2), trnI-GAU*(2), trnK-UUU*, trnL-CAA(2), trnL-UAA*, trnL-UAG(2), trnM-CAU, trnN-GUU(2), trnP-UGG, trnQ-UUG, trnR-ACG(2), trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC(2), trnV-UAC*, trnW-CCA, trnY-GUA, trnfM-CAU*	30(9)
Other genes	Maturase	*matK*	1
	Protease	*clpP***	1
	Envelope membrane protein	*cemA*	1
	Acetyl-CoA carboxylase	*accD***	1
	c-type cytochrome synthesis gene	*ccsA(2)*	1(1)
	Translation initiation factor	*infA*	1
	Conserved hypothetical chloroplast ORF	*ycf1, ycf15(2), ycf2(2), ycf3**, ycf4*	5(2)
Total			**133(19)**

(×2) indicates genes duplicated in the IR regions; *The genes containing a single intron; **The genes containing two introns

Table 3. SSRs identified in the chloroplast genomes of *P. obtusifolia*.

SSR type	Unit	Amount	Number	Ratio%
Mononucleotide	A/T	26	26	29.89%
Dinucleotide	AA/TT	26	30	34.48%
	AC/GT	1		
	AT/AT	3		
Trinucleotide	AAA/TTT	8	13	14.94%
	AAC/GTT	1		
	AAG/CTT	3		
	AAT/ATT	1		
Tetranucleotide	AAAA/TTTT	8	16	18.39%
	AAAC/GTTT	1		
	AAAG/CTTT	2		
	AAAT/ATTT	2		
	AACG/CGTT	1		
	AAGT/ACTT	1		
	ATAT/ATAT	1		
Pentanucleotide	AAAAG/CTTTT	2	2	2.30%
Total	——	——	87	——

Analysis of phylogenetics: To reveal phylogenetic position of *P. obtusifolia* with other members in *Premma*, phylogenetic analysis was performed based on *P. obtusifolia* protein sequence, and one taxon, *Plantago depressa* from Plantaginaceae was served as outgroup. MEGA7 was used to build maximal likelihood (ML) tree with 1,000 bootstrap (Kumar *et al*., 2016). There was the closest genetic relative between *P. obtusifolia* and *P. vietnamensis* (MT473774), closer genetic relative between *P. obtusifolia* (PP103654) and *P. szemaoensis* (MT473775), and the farthest kinship relationship between *P. obtusifolia* and *P. puberula* (NC_061379), *P. microphylla* (NC_026291) (Fig. 4). Cp genome sequence of *P. obtusifolia* in the investigation could provide helpful scientific basis for research in *Lamiaceae* plant.

Table 4. Repeat sequences in the *P. obtusifolia* cp genome.

Repeat type	Start position	End position	Size(bp)	Number	Ratio%
Forward	73	91867	91794	7	41.18%
	73	146142	146069		
	55	91885	91830		
	55	146160	146105		
	41	139264	139223		
	37	91903	91866		
	37	146178	146141		
Palindromic	25678	127737	102059	10	58.82%
	73	146124	146051		
	73	146142	146069		
	55	146124	146069		
	55	146160	146105		
	41	118190	118149		
	37	146124	146087		
	37	146178	146141		
	32	94	62		
	30	46243	46213		
Total	——	——	——	17	——

Table 5. The relative synonymous codon usage (RSCU) in *P. obtusifolia*.

Amino acid	Codon	Number	RSCU	Amino acid	Codon	Number	RSCU
Ter	UGA	13	0.75	Met	AUG	478	1.00
Ter	UAA	27	1.56	Asn	AAU	779	1.56
Ter	UAG	12	0.69	Asn	AAC	217	0.44
Ala	GCC	212	0.72	Pro	CCC	186	0.84
Ala	GCU	504	1.71	Pro	CCG	120	0.54
Ala	GCA	323	1.10	Pro	CCA	244	1.10
Ala	GCG	137	0.47	Pro	CCU	336	1.52
Cys	UGU	165	1.46	Gln	CAG	182	0.47
Cys	UGC	61	0.54	Gln	CAA	598	1.53
Asp	GAC	162	0.38	Arg	CGC	99	0.47
Asp	GAU	685	1.62	Arg	AGG	139	0.65
Glu	GAA	858	1.53	Arg	AGA	386	1.82
Glu	GAG	267	0.47	Arg	CGU	267	1.26
Phe	UUC	365	0.63	Arg	CGA	278	1.31
Phe	UUU	787	1.37	Arg	CGG	105	0.49
Gly	GGC	173	0.48	Ser	AGC	87	0.33
Gly	GGU	443	1.23	Ser	UCG	149	0.57
Gly	GGG	276	0.76	Ser	UCA	307	1.17
Gly	GGA	553	1.53	Ser	UCU	441	1.67
His	CAC	116	0.47	Ser	AGU	337	1.28
His	CAU	381	1.53	Ser	UCC	259	0.98
Ile	AUC	361	0.62	Thr	ACC	207	0.79
Ile	AUA	514	0.88	Thr	ACU	419	1.60
Ile	AUU	877	1.50	Thr	ACG	105	0.40
Lys	AAG	274	0.49	Thr	ACA	319	1.22
Lys	AAA	842	1.51	Val	GUU	443	1.55
Leu	UUA	697	1.90	Val	GUA	431	1.51
Leu	UUG	460	1.25	Val	GUG	150	0.52
Leu	CUC	129	0.35	Val	GUC	119	0.42
Leu	CUU	469	1.28	Trp	UGG	377	1.00
Leu	CUA	303	0.83	Tyr	UAC	137	0.36
Leu	CUG	143	0.39	Tyr	UAU	616	1.64

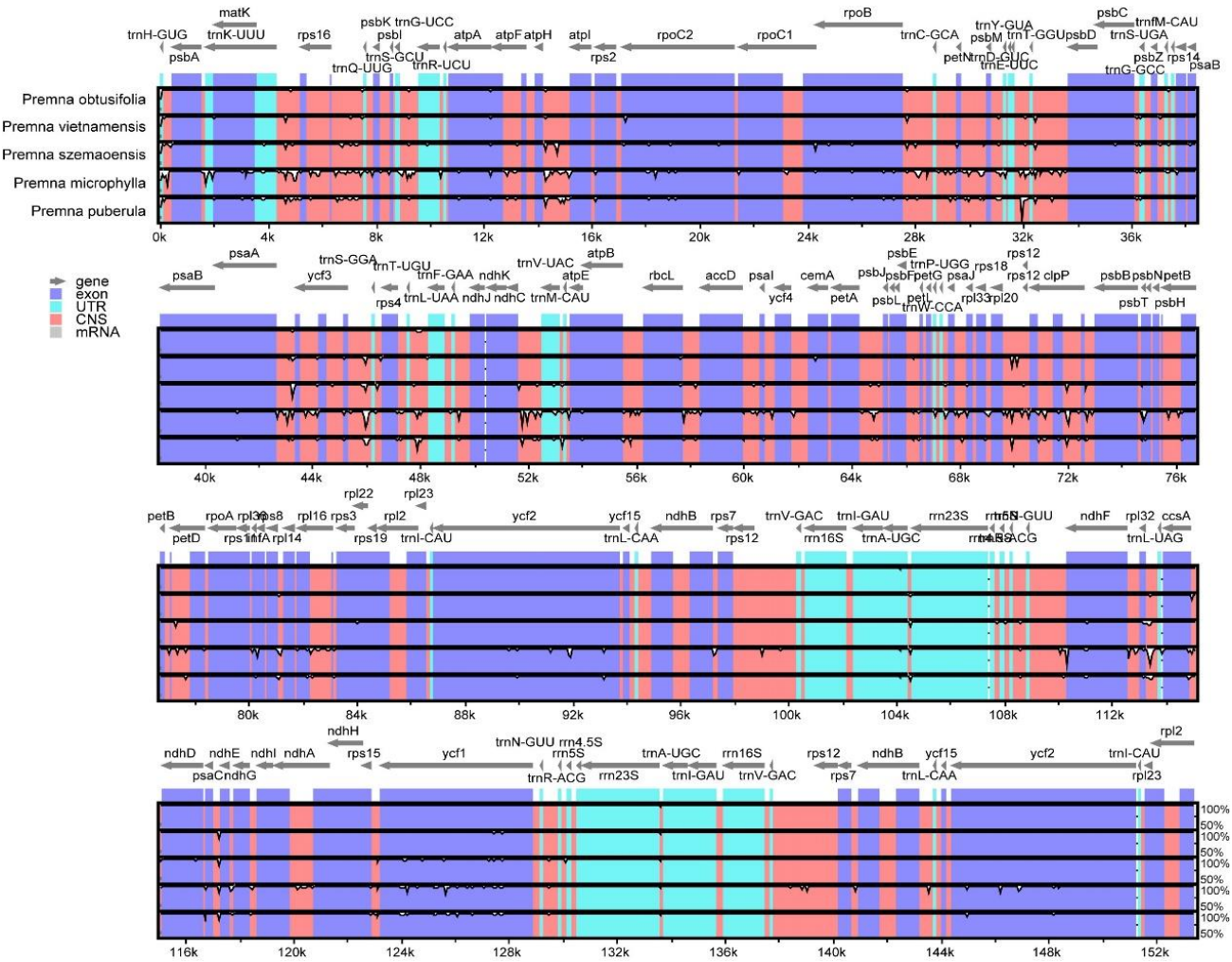

Fig. 2. Complete chloroplast genome comparison of five *Premna* accessions using the chloroplast genome of *P. obtusifolia* (PP103654) as a reference. Gray arrows and thick black lines above the alignment indicate gene orientation. Purple bars represent exons, sky-blue bars represent transfer RNA (tRNA) and ribosomal RNA (rRNA), red bars represent non-coding sequences (CNS), and white peaks represent differences of chloroplast genomes.

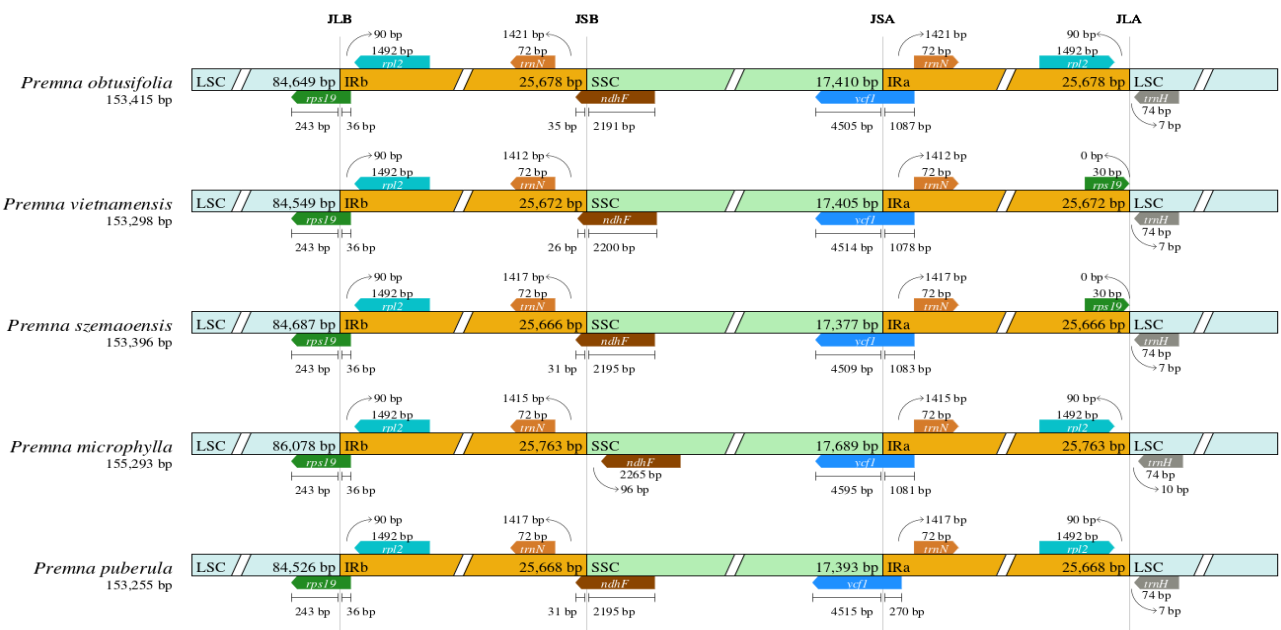

Fig. 3. Comparison of the borders of the LSC, SSC, and IR regions among four *Premma* cp genomes. Boxes above the main line indicate the adjacent border genes. The figure is not to scale with respect to sequence length, and only shows relative changes at or near the IR/SC borders.

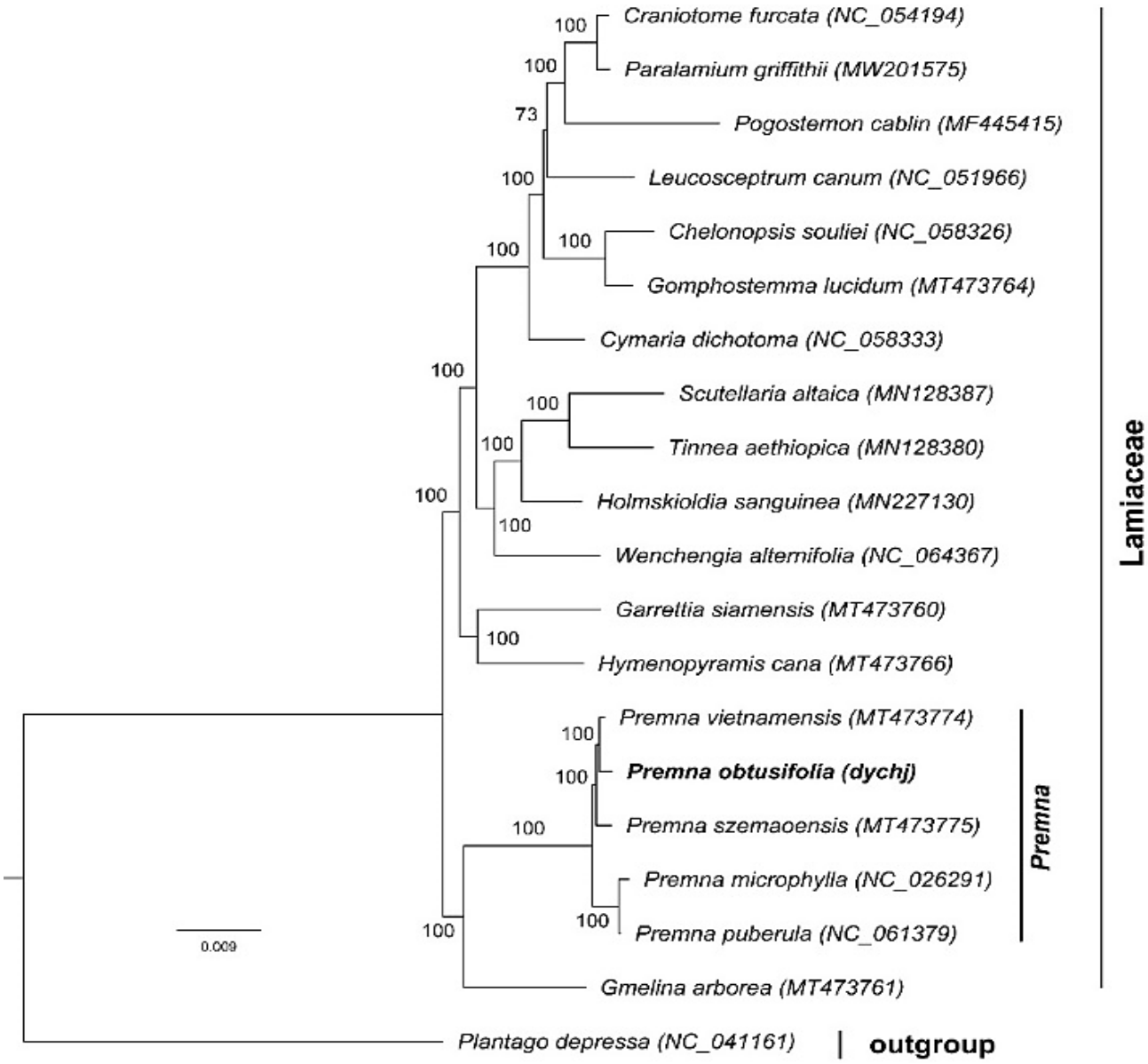

Fig. 4. Maximum likelihood phylogenetic tree for *P. obtusifolia* based on 20 complete chloroplast genomes

Discussion

P. obtusifolia, as one of semi-mangrove plants, plays an important role in biodiversity conservation of the bay estuary. This study assembled a complete and circular chloroplast genome with total length 153,415 bp, which exhibited higher GC content (43.21%) in IR region than those of the LSC (36.03%) and SSC (32.38%). The structure of *P. obtusifolia* cp genome was very conservative and relatively steady (Drouin *et al*., 2008), which showed very high sequence similarities (Wu *et al*., 2018) and close phylogenetic relationship among the same genus, such as *P. vietnamensis* (MT473774), *P. szemaoensis* (MT473775), *P. puberula* (NC_061379) and *P. microphylla* (NC_026291). Because of its small size of genome, slow evolution rate and little recombination, it is a good material to study species' evolution and classification (Ando *et al*., 2018). In order to better understand phylogenetic position and relationships of *P. obtusifolia* within family, more cp genomes from other genera will be needed in the future.

Conclusions

In the investigation, cp genome of *P. obtusifolia* was analyzed. It is forecasted to encode 114 unique genes, including 80 protein-coding genes, 30 tRNA genes, and 4 rRNA genes, with 19 duplicate genes in IR regions. 87 SSRs and 17 long repeats were recognized totally, and they may be employed as latent molecular markers. According to protein-coding genes in *P. obtusifolia* genome, codon usage indicated very low preference. The result suggests that codons displaying preference indicates the relative conservation of *P. obtusifolia*. The comparison showed that the category, quantity, and arrangement order of genes encoded by cp genome sequence of species in *Premna* are highly consistent; The variation between sequences primarily occurs in non-coding intergenic region.

Phylogenetic tree construction results strongly supported that *P. obtusifolia* was a close branch to *P. vietnamensis*. Those results indicate the species of *Premna*, with very conservative structure of cp genome, a good material for studying the protection of biological resources of *Premna* genus.

Acknowledgment

This study was supported by Science Research Project of Fuyang Normal University (2022KYQD0022), National Natural Science Foundation of China (NSFC, 31860085), Key Projects of Natural Science Research in Anhui Colleges and Universities (2024AH051451), Quality Engineering Projects of Graduate Education in Anhui (2023lhpysfjd053, 2024zyxwjxalk179), Biological and Medical Sciences of Applied Summit Nurturing Disciplines in Anhui Province (Anhui Education Secretary Department [2023]13). The funders have no specific role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

References

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol*., 215(3): 403-410.

Bankevich, A., S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V. M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, V. Nikolay, G. Tesler, M.A. Alekseyev and P.A. Pevzner. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comp. Biol*., 19(5): 455-477.

Beier, S., T. Thiel, T. Münch, U. Scholz and M. Mascher. 2017. MISA-web: a web server for microsatellite prediction. *Bioinformatics*, 33(16): 2583.

Chen, H.M., J.J. Shao, H. Zhang, M. Jiang, L.F. Huang, Z. Zhang, D. Yang, M. He, M. Ronaghi, X. Lou, B. Sun, W.W. Wu and C. Liu. 2018. Sequencing and analysis of *Strobilanthes cusia* (Nees) kuntze chloroplast genome revealed the rare simultaneous contraction and expansion of the inverted repeat region in angiosperm. *Front. Plant Sci*., 9: 324.

Curci P. L., D. D. Paola, D. Danzi, G. G. Vendramin and G. Sonnante. 2015. Complete chloroplast genome of the multifunctional crop globe artichoke and comparison with other Asteraceae. *PLoS One*, 10(3): e0120589.

Daniell, H., C. S. Lin, M. Yu and W. J. Chang. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol*., 17: 134.

Drouin G., H. Daoud and J. Xia. 2008. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol*., 49(3): 827-831.

Frazer K. A., L. Pachter, A. Poliakov, E. M. Rubin and I. Dubchak. 2004. VISTA: computational tools for comparative genomics. *Nucl. Acids Res*., 32: W273-W279.

Gao L., X. Yi, Y. X. Yang, Y. J. Su and T. Wang. 2009. Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. *BMC Evol. Biol*., 9: 130.

He S. L., Y. Yang, Z. W. Li, X. J. Wang, Y. B. Guo and H. Z. Wu. 2020. Comparative analysis of four *Zantedeschia* chloroplast genomes: expansion and contraction of the IR region, phylogenetic analyses and SSR genetic diversity assessment. *PeerJ*, 8: e9132.

Hu Y. H., K. E. Woeste and P. Zhao. 2017. Completion of the Chloroplast Genomes of Five Chinese Juglans and Their Contribution to Chloroplast Phylogeny. *Front. Plant Sci*., 7: 1955.

Huang H., C. Shi, Y. Liu, S. Y. Mao and L. Z. Gao. 2014. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships. *BMC Evol. Biol*., 14(1): 151.

Kanehisa M., S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki and M. Hirakawa. 2006. From genomics to chemical genomics: new developments in KEGG. *Nucl. Acids Res*., 34: D354-D357.

Kanehisa M., S. Goto, S. Kawashima, Y. Okuno and M. Hattori. 2004. The KEGG resource for deciphering the genome. *Nucl. Acids Res*., 32: D277-D280.

Kang L., D. F. Xie, Q. Y. Xiao, C. Peng, Y. Yu and X. J. He. 2019. Sequencing and analyses on chloroplast genomes of *Tetrataenium candicans* and two allies give new insights on structural variants, DNA barcoding and phylogeny in Apiaceae subfamily Apioideae. *Peer J*, 7(11): e8063.

Katoh K., K. Misawa, K. Kuma and T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucl. Acids Res*., 30(14): 3059-3066.

Kumar S., G. Stecher and K. Tamura. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol*., 33(7): 1870-1874.

Kurtz S. and C. Schleiermacher. 1999. REPuter: fast computation of maximal repeats in complete genomes. *Bioinformatics*, 15(5): 426-427.

Liu Q. P. and Q. Z. Xue. 2005. Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. *J. Genet.*, 84(1): 55-62.

Lohse M., O. Drechsel and R. Bock. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet*., 52: 267-274.

Lowe T. M. and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl. Acids Res*., 25(5): 955-964.

Mo Z., W. Lou, Y. Chen, X. Jia, M. Zhai, Z. Guo and J. Xuan. 2020. The chloroplast genome of *Carya illinoinensis*: genome structure, adaptive evolution, and phylogenetic analysis. *Forests*, 11(2): 207.

Nazareno A. G., M. Carlsen and L. G. Lohmann. 2015. Complete Chloroplast Genome of *Tanaecium tetragonolobum*: The First Bignoniaceae Plastome. *PLoS One*, 10(6): e0129930.

Ohsako T. and O. Ohmi. 2001. Nucleotide sequence variation of the chloroplast *trnK/matK* region in two wild *Fagopyrum* (Polygonaceae) species, *F. leptopodum* and *F. statice. Genes Genet. Syst*., 76(1): 39-46.

Pervaiz T., X. Sun, Y. Y. Zhang, R. Tao, J. H. Zhang and J. G. Fang. 2015. Association between chloroplast and mitochondrial DNA sequences in Chinese *Prunus* genotypes (*Prunus persica, Prunus domestica,* and *Prunus avium*). *BMC Plant Biol*., 15: 4.

Redwan R. M., A. Saidin and S.V. Kumar. 2015. Complete chloroplast genome sequence of MD-2 pineapple and its comparative analysis among nine other plants from the subclass Commelinidae. *BMC Plant Biol*., 15: 196.

Sablok G., K. C. Nayak, F. Vazquez and T. V. Tatarinova. 2011. Synonymous codon usage, GC(3), and evolutionary patterns across plastomes of three pooid model species: emerging grass genome models for monocots. *Mol. Biotechnol*., 49(2): 116-128.

Sharp P. M. and W. H. Li. 1987. The codon Adaptation Index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucl. Acids Res*., 15(3): 1281-1295.

Walker B. J., T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C. A. Cuomo, Q. D. Zeng, J. Wortman, S. K. Young and A. M. Earl. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, 9(11): e112963.

Wang W. B., H. Yu, J. H. Wang, W. J. Lei, J. H. Gao, X. P. Qiu and J. S. Wang. 2017. The complete chloroplast genome sequences of the medicinal plant *Forsythia suspensa* (Oleaceae). *Int. J. Mol. Sci*., 18(11): 2288.

Wang, W. Q., Q. Q. Li, Y. F. Chen, L. Q. Luo, L. T. Zhang and T. T. Li. 2016. Construction and evaluation of mangrove plant resource database in China [Review]. Xiamen: Xiamen University.

Wu M. L., Q. Li, J. Xu and X. W. Li. 2018. Complete chloroplast genome of the medicinal plant *Amomum compactum*: gene organization, comparative analysis, and phylogenetic relationships within Zingiberales. *Chin. Med*., 13: 10.

Wyman S., R. K. Jansen and J. L. Boore. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*, 20(17): 3252-3255.

Xu C., X. N. Cai, Q. Z. Chen, H. X. Zhou, Y. Cai and A. L. Ben. 2011. Factors affecting synonymous codon usage bias in chloroplast genome of oncidium gower ramsey. *Evol. Bioinform*., 2011(7): 271-278.

Yang J. B., M. Tang, H. T. Li, Z. R. Zhang and D. Z. Li. 2013. Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evol. Biol*., 13(1): 84.

Yi X., L. Gao, B. Wang, Y. J. Su and T. Wang. 2013. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): evolutionary comparison of cephalotaxus chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms. *Genome Biol. Evol*., 5(4): 688-698.

Young R. E. and S. Purton. 2016. Codon reassignment to facilitate genetic engineering and biocontainment in the chloroplast of *Chlamydomonas reinhardtii*. *Plant Biotechnol. J.*, 14(5): 1251-1260.

Yu X. Y, L. H. Zuo, D. D. Lu, B. Lu, M. S. Yang and J. M. Wang. 2018. Comparative analysis of chloroplast genomes of five *Robinia* species: Genome comparative and evolution analysis. *Gene*, 689: 141-151.

Zhang W,. Y. L. Zhao, G. Y. Yang, J. Peng, S. W. Chen and Z. G. Xu. 2019. Determination of the evolutionary pressure on *Camellia oleifera* on Hainan Island using the complete chloroplast genome sequence. *Peer J.*, 7: e7210.

Zuo L. H., A. Q. Shang, S. Zhang, X. Y. Yu, Y. C. Ren, M. S. Yang and J. M. Wang. 2017. The first complete chloroplast genome sequences of *Ulmus* species by de novo sequencing: Genome comparative and taxonomic position analysis. *PLoS One*, 12(2): e0171264.

(Received for publication 07 October 2024)