

## Research Article

# Chloroplast Genome Sequences and Comparative Analyses of Combretaceae Mangroves with Related Species

Ying Zhang, Hai-Li Li, Jun-Di Zhong, Yun Wang, and Chang-Chun Yuan 

Life Science and Technology School, Lingnan Normal University, Zhanjiang 524048, China

Correspondence should be addressed to Chang-Chun Yuan; [yuanchangchun@163.com](mailto:yuanchangchun@163.com)

Received 5 December 2019; Revised 1 September 2020; Accepted 12 September 2020; Published 30 September 2020

Academic Editor: Sankar Subramanian

Copyright © 2020 Ying Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the Combretaceae family, only two species of *Lumnitzera* and one species of *Laguncularia* belong to mangroves. Among them, *Lumnitzera littorea* (Jack) Voigt. is an endangered mangrove plant in China for the limited occurrence and seed abortion. In contrast, *Lumnitzera racemosa* Willd. is known as the most widespread mangrove plant in China. *Laguncularia racemosa* C. F. Gaertn., an exotic mangrove in China, has the fast growth and high adaptation ability. To better understand the phylogenetic positions of these mangroves in Combretaceae and in Myrtales and to provide information for studies on evolutionary adaptation for intertidal habitat, the complete chloroplast (cp) genomes of *Lu. racemosa* and *La. racemosa* were sequenced. Furthermore, we present here the results from the assembly and annotation of the two cp genomes, which were further subjected to the comparative analysis with *Lu. littorea* cp genomes we published before and other eleven closely related species within Myrtales. The chloroplast genomes of the three Combretaceae mangrove species: *Lu. littorea*, *Lu. racemosa*, and *La. racemosa* are 159,687 bp, 159,473 bp, and 158,311 bp in size. All three cp genomes host 130 genes including 85 protein-coding genes, 37 tRNAs, and 4 rRNAs. A comparative analysis of those three genomes revealed the high similarity of genes in coding-regions and conserved gene order in the IR and LSC/SSC regions. The differences between *Lumnitzera* and *Laguncularia* cp genomes are the locations of *rps19* and *rpl2* genes in the IR/SC boundary regions. Investigating the effects of selection events on shared protein-coding genes showed a relaxed selection had acted on the *ycf2*, *ycf1*, and *matK* genes of Combretaceae mangroves compared to the nonmangrove species *Eucalyptus aromaphloia*. The phylogenetic analysis based on the whole chloroplast genome sequence with one outgroup species strongly supported three Combretaceae mangroves together with other two Combretaceae species formed a cluster in Combretaceae. This study is the first report on the comparative analysis of three Combretaceae mangrove chloroplast genomes, which will provide the significant information for understanding photosynthesis and evolution in Combretaceae mangrove plants.

## 1. Introduction

Mangroves are critical marine resources for their remarkable ability to tolerate seawater and are uniquely adapted to tropical and subtropical coasts [1]. Mangrove forests played an important role in ecosystem services and supported coastal livelihoods, yet they are relatively low in the number of species [2]. In the world, the increasing number of endangered mangrove species has made a huge destruction on the global economy and environment [3, 4]. In the family Combretaceae, only two genera are typical mangrove constituents: *Lumnitzera* and *Laguncularia*. *Lumnitzera*, including 2 species *Lumnitzera littorea* and *Lumnitzera racemosa*, is an

Asian genus most characteristic of back mangal and ranges from India and Sri Lanka, through the Malasian region to China [5]. Due to natural and human impacts, populations of the two species in this genus have been isolated, fragmented, and highly disturbed [6]. *Laguncularia* is a monotypic genus and locally dominant in the West African and Caribbean regions of tropical America [5]. In China, two *Lumnitzera* species are both native mangrove species. *Lu. racemosa* can be found in almost all mangrove locations with the extensive adaptation [7], but *Lu. littorea* occurs only on Hainan Island and has become an endangered mangrove species due to its small population size caused by seed abortion [8, 9]. The only specie *La. racemosa* from the other genus was

firstly transplanted in Dongzhai harbor, Hainan, China [10]. Having the fast growing and high adaptation ability, it was used as the pioneer specie for the mangrove restoration in estuarine and coastal regions [11]. However, its invasiveness and the possibility of replacing native mangrove species have been subjects of debate in China and there are disagreements on whether *La. racemosa* should be planted [12]. Despite the obviously different adaptive capacities to the environment of those three mangrove species, *Lumnitzera* and *Laguncularia* are considered to be closely related and could have evolved from a common ancestor [5, 13].

In plants, chloroplast plays an important role in many cell functions, including photosynthesis, carbon fixation, and stress response [14], and is one of three major genetic systems [15]. A large amount of genetic information can get from the chloroplast genome to explore the occurrence, development, evolution of species, and to develop the fields of plant genomics and bioinformatics due to its self-replication mechanism and relatively independent evolution [16]. In angiosperms, the chloroplast genome was found to be a conserved quadripartite structure composing of two copies of inverted repeat (IR), one large single copy (LSC), and one small single copy (SSC) [14]. The chloroplast genome includes 120-130 genes, primarily participating in photosynthesis, transcription, and translation [17]. Recent studies have identified considerable diversity within noncoding intergenetic spacer regions, which often include important regulatory sequences [18]. Similar to the genes, the introns in chloroplast genomes of land plants are generally conserved, but the loss of introns within protein-coding genes and tRNA genes has been reported in several plant species [19]. Until now, over 3,665 plants have been sequenced, which has greatly improved chloroplast (cp) genome research (<https://www.ncbi.nlm.nih.gov/genome/browse#!/plasmids/>). However, there is little genetic and genomic research on plants from the family Combretaceae. As a large plant group, Combretaceae comprises over 600 species, but only few reports on sequencing cp genomes from this family. The whole cp information will help us to understand and progress the species identification, evolution, and genetic engineering in Combretaceae plants. Furthermore, only few of about 70 mangrove species have cp genome information, although the chloroplasts may play an important role for mangroves to inhabit the sea habitat [20]. Under salt stress, the obvious changes were found in the chloroplast ultrastructure, the expression of chloroplast genes, and even chloroplast proteins [21, 22].

In this study, two whole cp genomes of Combretaceae mangroves, *Lu. racemosa* and *La. racemosa*, were sequenced by using next-generation sequencing and applying a combination of de novo and reference guided assembly (*Lu. littorea* (MH551146) as a reference). Using a comparative genomics approach, we analyzed the characteristics of the cp genomes of the three mangrove species from Combretaceae. Nonsynonymous (Ka) and synonymous (Ks) substitution rates of conservative protein-coding genes among the three Combretaceae mangrove cp genomes were calculated to evaluate selection pressures. The coding sequences (CDSs) under selective events were also detected. Furthermore, the phylo-

genetic positions of three Combretaceae mangroves in Myrtales formed a clade sister to Lythraceae and Onagraceae, which will help to understand the role of natural selection in the adaptation of Combretaceae mangrove species.

## 2. Materials and Methods

**2.1. Plant Materials, DNA Extraction, and Sequencing.** Two Combretaceae mangrove samples were collected from Tielu Bay, Sanya, China (18°17'N, 109°44'E). Cp DNA was isolated from fresh leaf tissue of one individual plant of *Lu. racemosa* and *La. racemosa* separately. DNA extraction was used by purelink genomic plant DNA purification kit (Thermo Fisher, China). The extracted DNA was sent to a sequencing company TGS (Shenzhen, China) and sequenced with 350 bp pair-end reads by using an Illumina Hiseq-2000 platform according to the standard protocol at TGS.

**2.2. De Novo Assembly, Gap Filling, and Genome Annotation.** Raw reads were first filtered to obtain the high-quality clean data by removing adaptor sequences and low-quality reads with Q value  $\leq 20$ . SOAP de novo 2.04 (<http://soap.denovo.html>) was used to perform the initial assembly and obtain the contig sequences. The software GapCloser 1.12 (<http://soap.genomics.org.cn/soapdenovo.html>) was used to fill the gaps in the frame sequences diagram. The whole framework maps of the cp genomes were obtained by using the reference genes of *Lu. littorea* (MH551146) with the methods of [16]. The cp genome sequences were annotated with CpGAVAS software with parameters that use default values [23]. The genes in the cp genome were annotated using the DOGMA program with a parameter of 40 for the identity of the coding protein and other parameters are the default values [24], and then manually corrected. Star/stop codons and intron/exon borders were edited manually after comparison with the reference. In addition, we used tRNAscan-SE v2.0 to verify the identified tRNA genes [25]. The circular cp genome map was drawn using OGDRAW [26]. Primers were designed (Table S1) to test for correct sequence assembly. PCR amplification was performed according to a previously reported procedure [27].

**2.3. Sequence Divergence Analysis.** The complete cp genomes of *Lu. racemosa* and *La. racemosa* were compared with *Lu. littorea* by using the mVISTA program in Shuffle-LAGAN mode [28]. We set *Lu. littorea* as the reference sequence for these comparisons. The borders between single-copy regions (LSC and SSC) and inverted repeats (IR) regions among three Combretaceae cp sequences were compared by using Geneious v11.0.4 software. The sequence divergences in protein-coding genes between the three Combretaceae species were evaluated by using MEGA 7 [29]. To estimate the nucleotide diversity (Pi) values of LSC, SSC, and IR regions of three cp genomes, a sliding window analysis was conducted by using DnaSP 6 [30]. The step size was set to 200 bp, the window length was 600 bp, and the Tamura 3-parameter (T92) model was selected to test the pairwise sequence divergences [31].

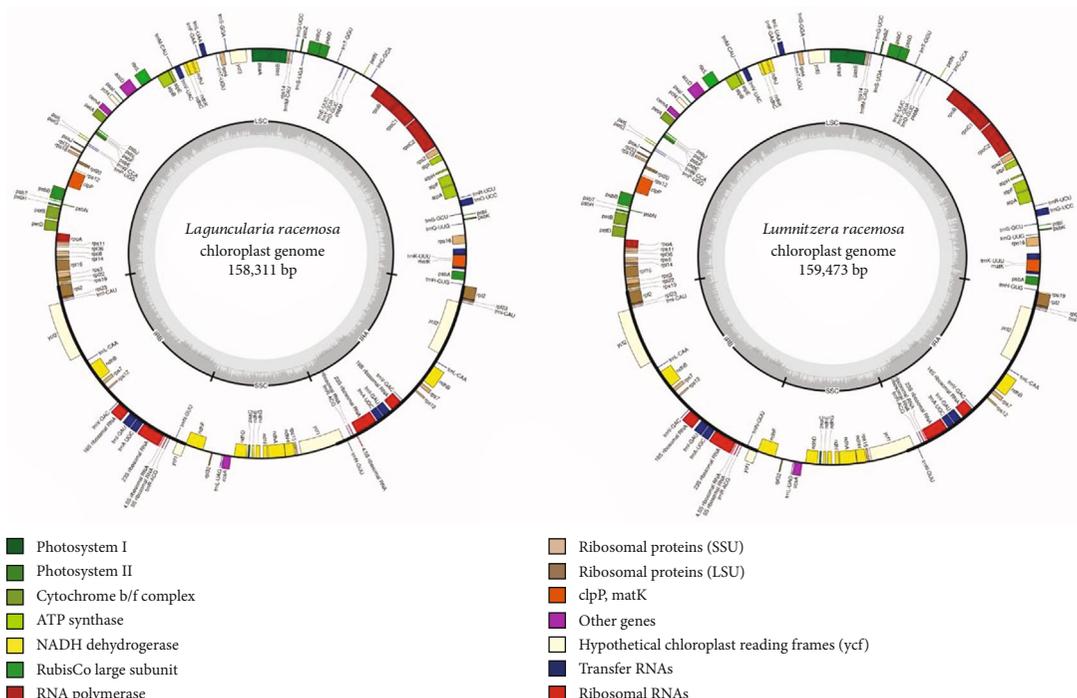


FIGURE 1: Gene maps of *Lu. racemosa* and *La. racemosa* chloroplast genomes. Genes drawn inside the circle are transcribed clockwise, and those outside are transcribed counterclockwise. Genes belonging to different functional groups are color-coded. The darker gray color in the inner circle corresponds to the GC content, and the lighter gray color corresponds to the AT content.

2.4. *Selection Pressure Analysis.* Selective pressures were analyzed for consensus protein-coding genes among four Myrtales species (*Lu. littorea*, *Lu. racemosa*, *La. racemosa*, and *Eucalyptus aromaphloia* (NC\_022396.1)). Seventy-six coding sequences (CDSs) longer than 300 bp were kept for the identification of codon usage patterns and then used for the estimation of codon usage using Codon W (<http://codonw.sourceforge.net>) [32]. Ka/Ks value for each gene was calculated by using the Ka/Ks calculator, and the settings were listed as previous description [32].

2.5. *Phylogenomic Analysis.* Phylogenomic analysis was performed for seventeen cp genomes including three Combretaceae mangroves in this study. Multiple sequence alignment was performed by using MAFFT software [33] to obtain the aligned chloroplast genomes before constructing the phylogenetic tree. The maximum likelihood tree, neighbor-joining tree, UPGMA tree, and test maximum parsimony tree were constructed using MEGA X [34] with *Rhizophora stylosa* as outgroup, and a bootstrap test was set to 1000 replicates to calculate each bootstrap value.

### 3. Results and Discussion

3.1. *Assembly and Features of cp Genomes for Three Species.* The complete cp genomes of *Lu. racemosa* and *La. racemosa* are 159,473 bp and 158,311 bp in length (GenBank accession number: MH551146, MH551145), respectively (Figure 1). The minor differences in length of cp genomes are no more than 214 bp in genus *Lumnitzera*. The maximum difference in length of cp genomes between the three Combretaceae

TABLE 1: Comparative analyses on the basic feature of the chloroplast genomes of three Combretaceae mangrove species.

	<i>Lu. littorea</i>	<i>Lu. racemosa</i>	<i>La. racemosa</i>
Length (bp)	159,687	159,473	158,311
GC content (%)	37.01	36.97	36.97
AT content (%)	62.99	63.03	63.03
LSC length (bp)	88,323	88,056	87,113
SSC length (bp)	18,558	18,613	18,886
IR length (bp)	26,403	26,402	25,156
Gene number	130	130	130
Gene number in IR regions	39	38	36
Protein-coding gene number	85	85	85
Protein-coding gene (%)	65.1	65.4	65.1
rRNA gene number	8	8	8
rRNA (%)	6.2	6.2	6.2
tRNA gene number	37	37	37
tRNA (%)	28.7	28.5	28.7

mangroves is 1,376 bp (Table 1). The typical quadripartite structure of most angiosperms was found in both of the two cp genomes, which comprise a pair of IRs (26,402 bp for *Lu. racemosa* and 26,156 bp for *La. racemosa*) separated by the LSC (88,056 bp for *Lu. racemosa* and 87,023 bp for *La. racemosa*) and SSC (18,613 bp for *Lu. racemosa* and 18,887 bp for *La. racemosa*) regions (Table 1). The cp

TABLE 2: Genes in the cp genomes of three Combretaceae mangrove species.

Category	Gene group	Gene name	
Photosynthesis	Photosystem I	<i>psaA, psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>	
	Photosystem II	<i>psaA, psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>	
	Cytochrome b/f complex	<i>petA, petB, petD, petG, petL, petN</i>	
	ATP synthesis	<i>atpA, atpB, atpE, atpF, atpH, atpI</i>	
	Large subunit of RuBisCo	<i>rbcL</i>	
	NADH dehydrogenase	<i>ndhA, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>	
	Ribosomal RNA genes	<i>rrn4.5, rrn5, rrn16, rrn23</i>	
	Ribosomal RNA genes (SSU)	<i>rps2, rps3, rps4, rps7, rps8, rps12 *, rps14, rps15, rps16, rps18, rps19</i>	
	Ribosomal RNA genes (LSU)	<i>rpl2, rpl14, rpl16, rpl20, rpl22, rpl23, rpl32, rpl33, rpl36</i>	
	RNA polymerase	<i>rpoA, rpoB, rpoC1, rpoC2</i>	
Self-replication	Transfer RNA genes	<i>trnA-UGC, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnM-CAU, trnG-UCC, trnH-GUG, trnI-CAU, trnI-GAU, trnK-UUU, trnL-CAA, trnL-UAA, trnL-UAG, trnM-CAU, trnN-GUU, trnP-UGG, trnQ-UUG, trnR-ACG, trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC, trnV-UAC, trnW-CCA, trnY-GUA</i>	
	Maturase	<i>matK</i>	
	Envelope membrane factor	<i>cemA</i>	
	Subunit of acetyl-CoA	<i>accD</i>	
	C-type cytochrome synthesis gene	<i>ccsA</i>	
	Protease	<i>clpP</i>	
	Hypothetical chloroplast reading	<i>ycf1, ycf2, ycf3, ycf4</i>	

\*Pseudogene.

genomes of the three Combretaceae mangroves are considerable conservation in length, and a previous study showed that the IR sequence length of some species in the Myrtales is 25.7-28.7 kb [35]. The same GC content in cp genomes (36.97%) was revealed for both *Lu. racemosa* and *La. racemosa*, similar to the value reported previously for *Lu. littorea* cp genome [36]. A total of 130 genes including 85 protein-coding genes, 37 tRNA genes, and 8 rRNA genes were found in three cp genomes. The only difference lies in the gene number in the IR regions between the three cp genomes (Table 1). The overall genomic structure including gene number and gene order were well-conserved.

Same as observed for other angiosperms, the protein-coding genes in the cp genomes of Combretaceae mangroves include five genes encoding photosystem I components (*psaA, B, C, I, and J*) and fifteen genes related to photosystem II (*psbA, B, C, D, E, F, H, I, J, K, L, M, N, T, and Z*), nine genes encoding large ribosomal proteins (*rpl2, 14, 16, 20, 22, 23, 32, 33, and 36*), eleven genes encoding small ribosomal proteins (*rps2, 3, 4, 7, 8, 12, 14, 15, 16, 18, and 19*), six genes encoding ATP synthase and electron transport chain components, and eleven genes encoding NADH dehydrogenase (Table 2). For the genes having two gene copies in IRs regions, *rps19* is found in the cp genome of *La. Racemosa* and same with *Lu. littorea* [36], but in *Lu. racemosa* has only one *rps19* gene. For the genes containing an intron, *trnK-UUU*, *trnV-UAC*, and *trnG-UCC* are found in cp genome of *La. Racemosa*, but in *Lu. racemosa* no intron is found in both of them. Other

gene *trnR-AGG* with one intron is found in the cp genome of *Lu. racemosa*, but no intron in the cp genome of *La. racemosa* (Tables S2 and S3). The majority of reported intron losses have been observed in specific plant groups or species including monocots, eudicots, and gymnosperms [17].

**3.2. Comparison of cp Genomes of Three Combretaceae Mangroves.** Gene content and structure are found to be conserved in the chloroplast genome sequences of most autotrophic land plants, but there are still some protein-encoding genes being absent in some specific species. *InfA*, *rpl22*, and *ndh* were the most reported genes are found from the chloroplast genomes to nuclear or mitochondrial genomes [17, 37]. But in this study, only gene *infA* was absent in the three cp genomes, and the same loss events are found in other Myrtales plants [38]. For the translation initiation factor gene, *infA* in cp has been lost independently at least 24 times in angiosperms and evidence provided from some cases suggested functional replacement by a nucleus copy [39]. The essential gene *rpl22*, which is reported in 57 chloroplast genomes to have been deleted from the chloroplast and transferred to the nuclear genome [17], is kept in three cp genomes of Combretaceae mangroves. The *ndh* proteins assemble into the photosystem I complex to mediate cyclic electron transport in chloroplasts and play an important role in facilitating chlororespiration, and in some autotrophic plants which were found to be lost in cp genomes [24]. In the three Combretaceae mangroves, all the eleven *ndh* genes



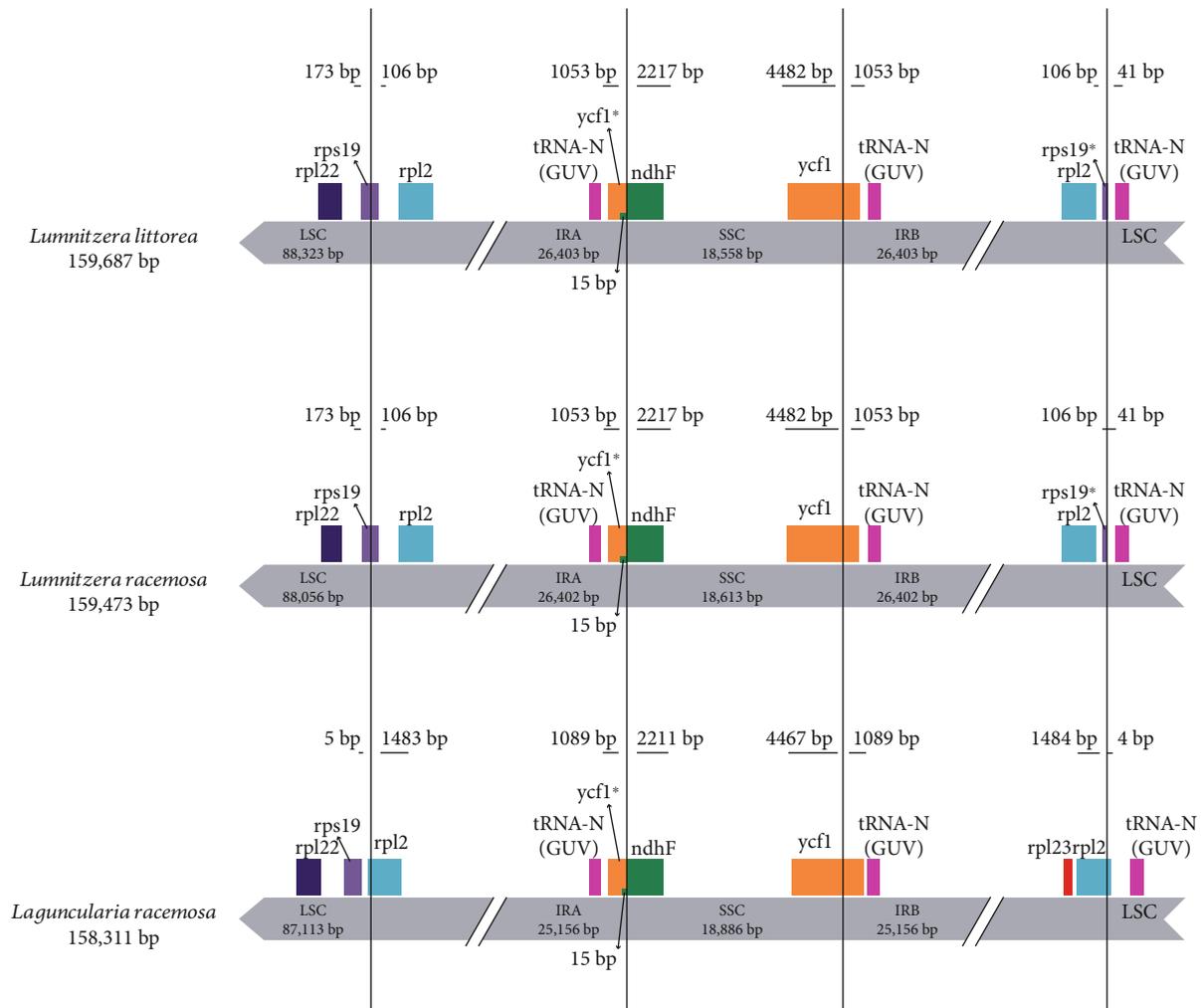


FIGURE 3: Comparison of IR, LSC, and SSC junction positions among the three chloroplast genomes. The features drawn are not to scale. The symbol \* means pseudogene created by IRb/SSC border extension into *ycf1* genes and IRb/LSC border extension into *rps19* genes.

crossed the IRB/LSC boundary creating *rps19* pseudogene fragment were only be found in *Lu. littorea* and *Lu. racemosa* cp genomes. However, in *La. racemosa*, the *rps19* gene was located in the LSC region, 5 bp apart from the IRb/LSC border. The *rpl2* genes crossed the LSC/IRB and LSC/IRA junction in *La. racemosa* cp genome but the distances between *rpl2* and the border is 106 bp in both of the two *Lumnitzera* species. Those variations at the IR/SC borders in these three cp genomes contribute to the differences in length of the cp genome sequence as a whole and were found in other plants [31].

**3.3. Genome Sequence Divergence among Combretaceae Species.** The sequence divergence among the three Combretaceae mangrove cp genomes was investigated by calculating the nucleotide variability (Pi) values within 600 bp windows (200 bp stepwise moving) in LSC, SSC, and IR regions (Figure 4). In the LSC, SSC, and IR regions, the values were found to vary from 0 to 0.22 with a mean of 0.085, from 0 to 0.095 with a mean of 0.063 and from 0 to 0.0169 with a mean of 0.082 separately. These results mean conservation

between the three Combretaceae mangrove cp genomes. However, the certain highly variable regions have also been found in *rpoC2* and *ndhC/trnV* with  $P_i > 0.02$  in LSC regions, *trnK/matK* with  $P_i > 0.08$  in SSC regions, and *rpoC1* and *rpoB* with  $P_i > 0.01$  in IRs regions (Figure 4). All those regions or some of them have also been identified as highly variable in other plants, which showed great potential as sources of useful phylogenetic markers for Combretaceae and other plant species [31, 45].

**3.4. Selection Events in Protein-Coding Genes.** The nonsynonymous (Ka) and synonymous (Ks) substitution ratios (Ka/Ks) were calculated for 78 consensus protein-coding genes to estimate selective pressures in cp genomes of the three Combretaceae mangroves and one reference species *E. aromaphloia*, which is a normal land and nonhalophytes plant, to evaluate the selective pressure (Table S4). The average Ka/Ks ratio of the shared genes analyzed across the three cp genomes was 0.1638. Although all of the Ka/Ks values were all less than 1.0 in Figure 5, the Ka/Ks ration of three genes (*ycf1*, *ycf2*, and *matK*) in all Combretaceae

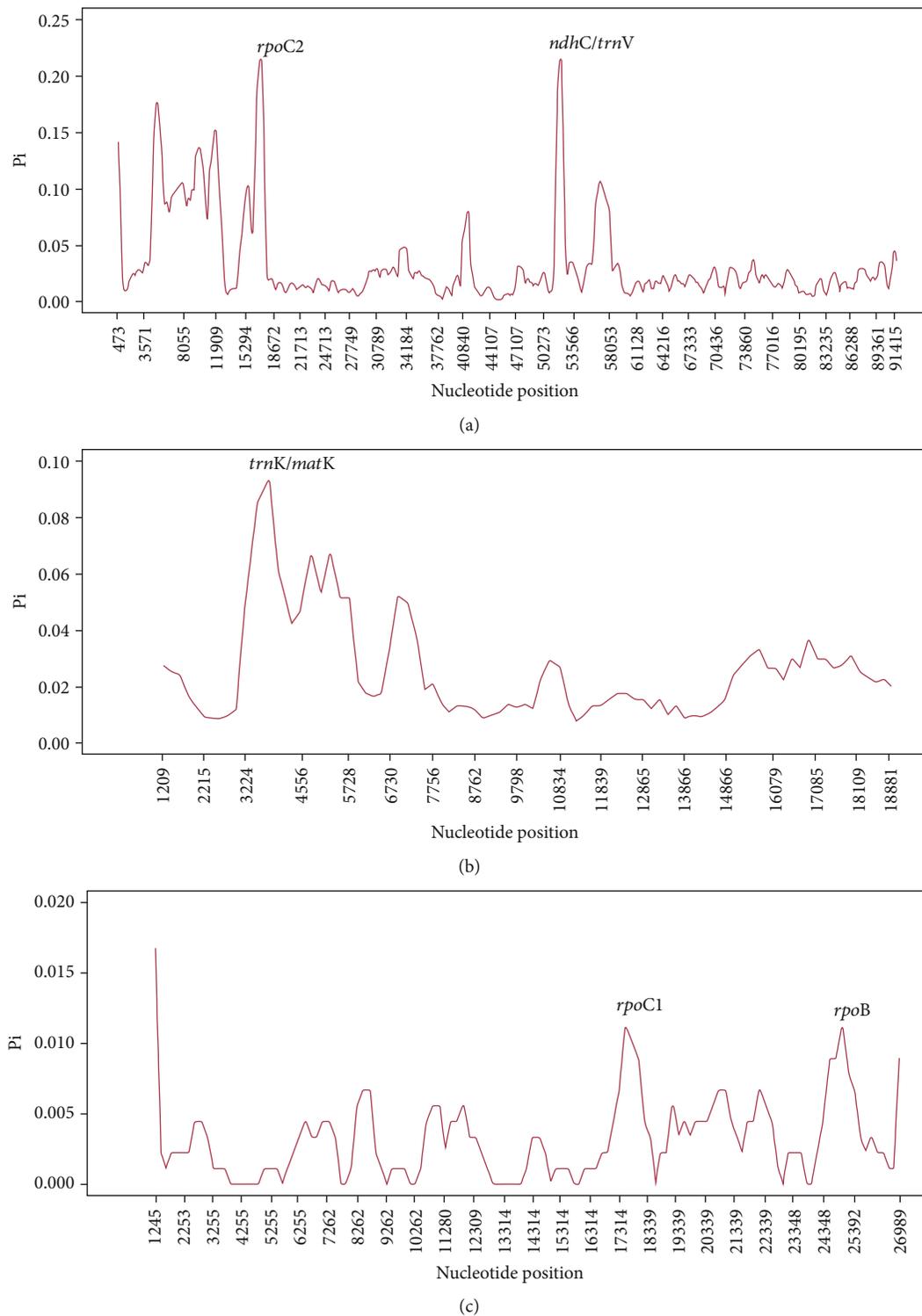


FIGURE 4: Comparative analysis of nucleotide variability (Pi) values among three Combretaceae mangrove cp genome sequences. (a) Analysis of the LSC regions; (b) Analysis of the SSC regions; (c) Analysis of the IR regions. (Window length: 600 bp, step size: 200 bp). x-axis: position of the midpoint of a window, y-axis: nucleotide diversity of each window.

mangroves were within the range of 0.5 to 1.0 indicating a relaxed selection [31]. Furthermore, there was one gene *rpl22* only found in *La. racemosa* with the Ka/Ks ratio within the range of 0.5 to 1.0. Another gene *accD* with the Ka/Ks ratio is closed to 0.5 in *La. racemosa* and not in the

other two mangroves in Figure 5. The most conserved genes with Ka/Ks values of 0 in the three cp genomes were *atpH*, *petG*, *petN*, *psaC*, *psbE*, *psbM*, *psbN*, *psbT*, and *rps7*, suggesting very strong purifying selection (Table S4). For the none positive selection gene was found in

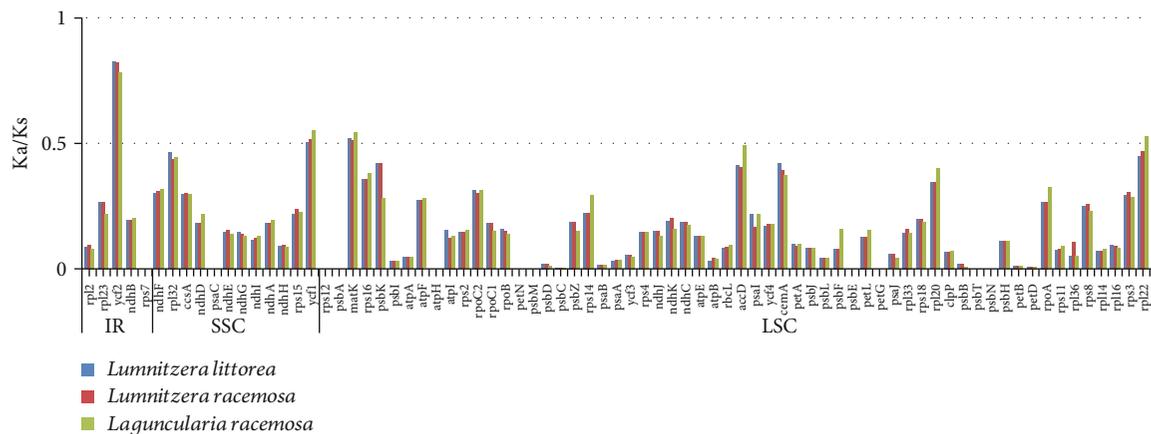


FIGURE 5: Ka/Ks ratios for protein-coding genes from *Lu. littorea*, *Lu. racemosa*, and *La. racemosa* chloroplast genome in comparison with *Eucalyptus aromaphloia*.

Combretaceae mangroves cp genome, this may be because adaptive modifications to salt stresses targeting genes in the nucleus were sufficient to maintain homeostasis for photosynthesis since there are a variety of strategies for plants to adapt to the environment, so there is no need for adaptive evolution of chloroplast-encoded genes [46]. But genes under relative selection in cp genomes may play some roles in Combretaceae mangroves evolution and environment adaptation. *Ycf* genes have proved useful for analyzing cp genome variation in higher plants and algae, even though the function is not thoroughly known [47]. Among *Ycf* genes, *ycf1* and *ycf2* are the two largest genes and are located in IR/SC junction and IR region, respectively, almost in all plants [48]. The gene *ycf1* encodes a component of the chloroplast's inner envelope membrane protein translocation [49]. It is also highly variable in terms of phylogenetic information at the level of species, has also been shown to be subject to relative selection with *rps12* and *matK* genes in three Combretaceae mangroves, and has also been identified in many plant lineages [50]. The *ycf2* gene in the cp genome is regarded as having one of the fastest evolutionary rates within the cp genome for the lost events [51]. Even though a giant reading frame of *ycf2* is believed as unknown function in land plants, which is responsible for the differences in the competitive behavior of plastid genotypes [52]. The nucleotide sequence similarity of *ycf2* in land plant is extraordinarily low compared to other plastid-encoded genes, being less than 50% across bryophytes, ferns, and seed plants [53]. The gene *matK* (maturase K gene) is a plant chloroplast gene [54] and encodes an intron maturase involved in the cutting and splicing of Group II RNA transcriptional introns [55]. By sequence analysis, there are one well-conserved domain X and the remnants of a reverse transcriptase domain in the coding sequence of *matK* [55]. The *matK* gene is always found to be located within the intron of the *trnK* gene (Figure 1) same with some land plants [56]. *MatK* gene is identified to be essential for plant cell survival, and the expression of it needs to be tightly regulated to prevent detrimental effects and establishes another link between leaf variegation and

chloroplast translation [57]. For the studies of plant systematics, the *matK* gene is the high selectively used gene as the DNA barcoding fragment for plants [58].

Gene *rpls* are ribosome genes that have been proven to be essential for the chloroplast ribosome development in plants [59]. Among them, *rpl22* is always found to have a frameshift mutation generated premature termination codons within it in land plants. In some species, *rpl22* was also found to be truncated with considerable length variation [60]. For the plastid gene, *accD*, which encodes the  $\beta$ -carboxyl transferase subunit of acetyl-CoA carboxylase, is an essential and required component for plant leaf development [61]. The same site-specific selection events of it have been observed in other plants [8], and which have been shown to affect the plant fitness by altering the acetyl-CoA carboxylase production [62]. Both genes *rpl22* and *accD* were only found to be under different selection between *La. racemosa* and two *Lumnitzera* species, and for *La. racemosa*, which is an introductive mangrove species in China and having the different original habitat other two *Lumnitzera* mangroves [5].

**3.5. Phylogenetic Analysis of Three Combretaceae Mangrove cp Genomes and Related Myrtale Species.** Cp genome sequences are useful for deciphering phylogenetic relationships among closely related taxa and for clarifying the evolutionary patterns of plant species [63]. To evaluate the phylogenetic position of Combretaceae mangrove species in the Myrtales, the whole chloroplast genome sequences of seven Combretaceae species with other nine plants in Myrtales were used to infer their phylogenetic relationships (Figure 6). Among those cp genomes, one mangrove plant *Rhizophora stylosa* was set as outgroup. By the phylogenetic analysis, 16 Myrtale species were clustered into five families, six species in Myrtaceae, two species in Onagraceae, and five Combretaceae mangroves were clustered into their own groups. Furthermore, two *Lumnitzera* genera were clustered into one subclade in the Combretaceae group. Within the family Combretaceae, *Laguncularia* and *Lumnitzera* are quite closely related in the field in China and are considered to be evolved from a common ancestor [5]. In this study, the position of Combretaceae mangroves are more closed with

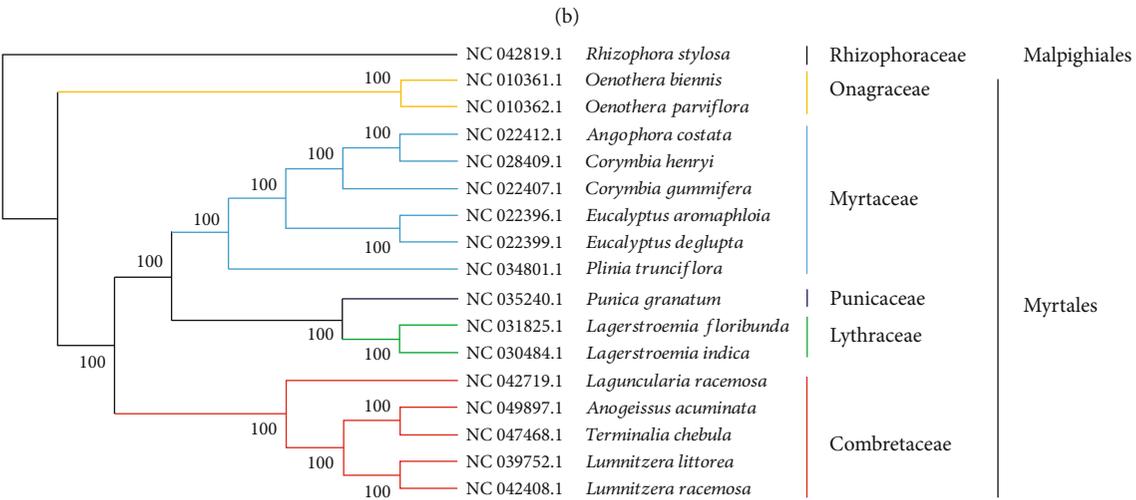
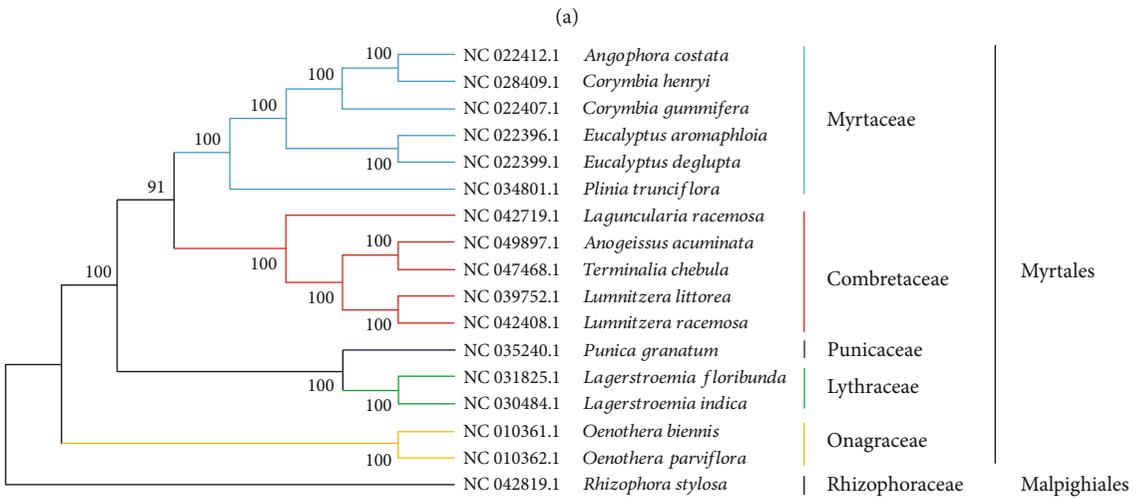
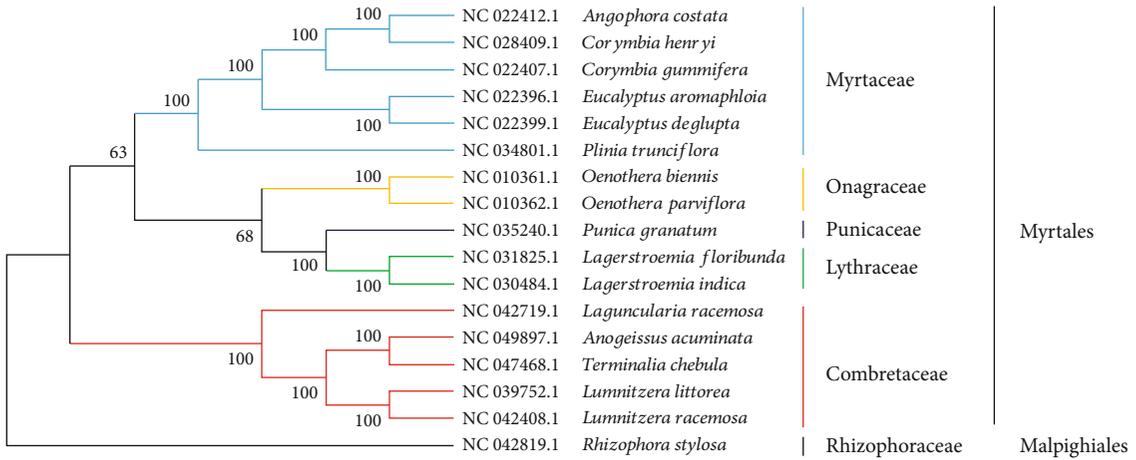


FIGURE 6: Continued.

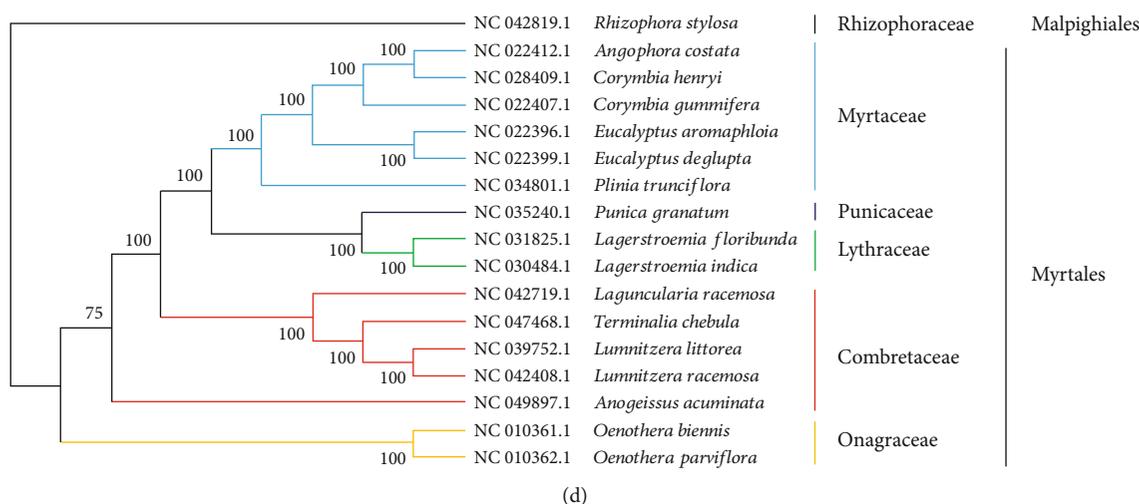


FIGURE 6: Phylogenetic relationship between five Combretaceae cp genomes and 11 related species in Myrtales, inferred from the whole cp genome sequences. Rhizophoraceae plant: *Rhizophora stylosa* as outgroup. (a) Maximum likelihood tree; (b) test maximum parsimony tree; (c) neighbor-joining tree; (d) UPGMA tree.

Lythraceae and Punicaceae species in Myrtales, which was showed in all the four polygenetic trees. In Myrtales, there included three significant shifts in diversification rates, one of them contributed from Combretaceae [64]. The chloroplast genome of those three Combretaceae mangroves will provide valuable and essential genetic information to further the phylogenetic resolution among angiosperms [63].

#### 4. Conclusions

With the help of high-throughput sequencing technology, we comparatively analyzed the complete cp genomes of three Combretaceae mangroves. The gene contents and gene orders of the three cp genomes were highly conserved. Gene *ycf1*, *rps19*, and *ndhF* were found to be the most divergent coding regions, and there are still some other noncoding regions with the high sequence divergence, which could potentially serve as molecular markers in phylogenetic studies. Three genes *ycf1*, *ycf2*, and *matK* were found to be under relaxed selection in the cp genomes of three Combretaceae mangroves. Phylogenetic analysis showed that the position of Combretaceae mangroves is closest to Punicaceae and Lythraceae species in Myrtales.

#### Data Availability

The analysis and experimental data used to support the findings of this study are included within the article.

#### Conflicts of Interest

The authors declare no conflict of interest.

#### Authors' Contributions

Y.Z. and C.-C.Y. analyzed the data and wrote the manuscript; J.-D.Z. and Y.W. conceived the experiments; J.-D.Z. per-

formed the experiments. All authors read and approved the final manuscript. Ying Zhang and Hai-Li Li are co-authors.

#### Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant Nos. 41776148 and 31760119).

#### Supplementary Materials

The following are available online at <https://www.mdpi.com/xxx/s1>. Table S1: primers used for gap closing and sequencing verification in *Laguncularia racemosa* and *Lumnitzera racemosa*. Table S2: functional classification of *Lumnitzera racemosa* chloroplast genome genes. Table S3: functional classification of *Laguncularia racemosa* chloroplast genome genes. Table S4: Ka/Ks value of the consensus protein-coding genes in three cp genomes of Combretaceae mangroves. (Supplementary Materials)

#### References

- [1] E. Yang, S. Yi, F. Bai et al., "Cloning, characterization and expression pattern analysis of a cytosolic copper/zinc superoxide dismutase (*SaCSD1*) in a highly salt tolerant mangrove (*Sonneratia alba*)," *International Journal of Molecular Sciences*, vol. 17, no. 1, p. 4, 2016.
- [2] B. A. Polidoro, K. E. Carpenter, L. Collins et al., "The loss of species: mangrove extinction risk and geographic areas of global concern," *PLoS One*, vol. 5, no. 4, article e10095, 2010.
- [3] N. Duke, J. Meynecke, S. Dittmann et al., "A world without mangroves?," *Science*, vol. 317, no. 5834, pp. 41-42, 2007.
- [4] B. H. Daru, K. Yessoufou, L. T. Mankga, and T. J. Davies, "A global trend towards the loss of evolutionarily unique species in mangrove ecosystems," *PLoS One*, vol. 8, no. 6, article e66686, 2013.
- [5] P. Tomlinson, *The Botany of Mangroves*, Cambridge University Press, Cambridge, UK, 1986.

- [6] G. Su, Y. Huang, F. Tan, X. Ni, T. Tang, and S. Shi, "Conservation genetics of *Lumnitzera littorea* (Combretaceae), an endangered mangrove, from the Indo-West Pacific," *Marine Biology*, vol. 150, no. 3, pp. 321–328, 2007.
- [7] J. Li, Y. Yang, Q. Chen et al., "Pronounced genetic differentiation and recent secondary contact in the mangrove tree *Lumnitzera racemosa* revealed by population genomic analyses," *Scientific Reports*, vol. 6, no. 1, article 29486, 2016.
- [8] Y. Zhang, C. R. Zhong, S. Li, T. L. Yan, and W. Guan, "Endangered species of mangrove plants: *Lumnitzera littorea*," *Forest Resources Management (China)*, vol. 5, pp. 103–107, 2013.
- [9] Y. Zhang, Y. Li, X. Zhang, and Y. Yang, "Flower phenology and breeding system of endangered mangrove *Lumnitzera littorea* (Jack.) Voigt," *Chinese Journal of Applied & Environmental Biology*, vol. 23, no. 1, pp. 77–81, 2017.
- [10] B. Liao, S. Zheng, Y. Chen, L. Mei, W.-J. Zeng, and D.-Z. Zheng, "Preliminary report on introduction of several alien mangrove plants in Dongzhai harbour of Hainan province," *Journal of Central South Forestry University*, vol. 26, no. 1, pp. 63–67, 2006.
- [11] W. H. Lin, C. A. Zhan, D. X. Zheng, and L. Li, "Study on afforestation of six mangrove species in sandy beach of eastern Guangdong," *Forestry Science and Technology of Guangdong Province*, vol. 30, no. 2, pp. 69–71, 2014.
- [12] X. L. Wang, C. Y. Lu, L. Zhou et al., "Allelopathic effects of exotic mangrove species *Languncularia racemosa* on *Bruguiera gymnorhiza*," *Journal of Xiamen University (Natural Science)*, vol. 56, no. 3, pp. 239–245, 2017.
- [13] F. Tan, S. Shi, Y. Zhong, X. Gong, and Y. Wang, "Phylogenetic relationships of Combretaceae (Combretaceae) inferred from plastid nuclear gene and spacer sequences," *Journal of Plant Research*, vol. 115, pp. 475–481, 2002.
- [14] F. L. Li, L. Yang, Q. J. Zan et al., "Does energetic cost for leaf construction in *Sonneratia* change after introduce to another mangrove wetland and differ from native mangrove plants in South China?," *Marine Pollution Bulletin*, vol. 124, no. 2, pp. 1071–1077, 2017.
- [15] L.-H. Zuo, A.-Q. Shang, S. Zhang et al., "The first complete chloroplast genome sequences of *Ulmus* species by de novo sequencing: genome comparative and taxonomic position analysis," *PLoS One*, vol. 12, no. 2, article e0171264, 2017.
- [16] P. Li, S. Zhang, F. Li et al., "A phylogenetic analysis of chloroplast genomes elucidates the relationships of the six economically important *Brassica* species comprising the triangle of U," *Frontiers in Plant Science*, vol. 8, p. 111, 2017.
- [17] H. Daniell, C.-S. Lin, M. Yu, and W.-J. Chang, "Chloroplast genomes: diversity, evolution, and applications in genetic engineering," *Genome Biology*, vol. 17, no. 1, p. 134, 2016.
- [18] H. Daniell, S.-B. Lee, J. Grevich et al., "Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other Solanaceae genomes," *Theoretical and Applied Genetics*, vol. 112, no. 8, pp. 1503–1518, 2006.
- [19] R. K. Jansen, Z. Cai, L. A. Raubeson et al., "Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns," *Proceedings of the National Academy of Sciences*, vol. 104, no. 49, pp. 19369–19374, 2007.
- [20] J. dos Santos Garcia, Â. C. Dalmolin, M. G. C. França, and P. A. O. Mangabeira, "Different salt concentrations induce alterations both in photosynthetic parameters and salt gland activity in leaves of the mangrove *Avicennia schaueriana*," *Ecotoxicology and Environmental Safety*, vol. 141, pp. 70–74, 2017.
- [21] F. Bejaoui, J. J. Salas, I. Nouairi et al., "Changes in chloroplast lipid contents and chloroplast ultrastructure in *Sulla carnosa* and *Sulla coronaria* leaves under salt stress," *Journal of Plant Physiology*, vol. 198, pp. 32–38, 2016.
- [22] P. Y. Yousuf, A. Ahmad, I. M. Aref et al., "Salt-stress-responsive chloroplast proteins in *Brassica juncea* genotypes with contrasting salt tolerance and their quantitative PCR analysis," *Protoplasma*, vol. 253, no. 6, pp. 1565–1575, 2016.
- [23] C. Liu, L. Shi, Y. Zhu et al., "CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences," *BMC Genomics*, vol. 13, no. 1, p. 715, 2012.
- [24] T. Wakasugi, J. Tsudzuki, S. Ito, K. Nakashima, T. Tsudzuki, and M. Sugiura, "Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*," *Proceedings of the National Academy of Sciences*, vol. 91, no. 21, pp. 9794–9798, 1994.
- [25] P. Schattner, A. N. Brooks, and T. M. Lowe, "The tRNAscan - SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs," *Nucleic Acids Research*, vol. 33, pp. 686–689, 2005.
- [26] M. Lohse, O. Drechsel, S. Kahlau, and R. Bock, "Organellar-GenomeDRAW—va suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets," *Nucleic Acids Research*, vol. 41, no. w1, pp. 575–581, 2013.
- [27] S. Asaf, M. Waqas, A. L. Khan et al., "The complete chloroplast genome of wild rice (*Oryza minuta*) and its comparison to related species," *Frontiers in Plant Science*, vol. 8, p. 304, 2017.
- [28] K. A. Frazer, L. Pachter, A. Poliakov, E. M. Rubin, and I. Dubchak, "VISTA: computational tools for comparative genomics," *Nucleic Acids Research*, vol. 32, pp. 273–279, 2004.
- [29] S. Kumar, G. Stecher, and K. Tamura, "MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets," *Molecular Biology and Evolution*, vol. 33, no. 7, pp. 1870–1874, 2016.
- [30] J. Rozas, A. Ferrer-Mata, J. C. Sánchez-DelBarrio et al., "DnaSP 6: DNA sequence polymorphism analysis of large data sets," *Molecular Biology and Evolution*, vol. 34, no. 12, pp. 3299–3302, 2017.
- [31] K. Yin, Y. Zhang, Y. Li, and F. K. Du, "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," *International Journal of Molecular Sciences*, vol. 19, no. 4, p. 1042, 2018.
- [32] Z. Ivanova, G. Sablok, E. Daskalova et al., "Chloroplast genome analysis of resurrection tertiary relict *Haberlea rhodopensis* highlights genes important for desiccation stress response," *Frontiers in Plant Science*, vol. 8, p. 204, 2017.
- [33] J. Rozewicki, S. Li, K. M. Amada, D. M. Standley, and K. Katoh, "MAFFT-DASH: integrated protein sequence and structural alignment," *Nucleic Acids Research*, vol. 47, pp. W5–W10, 2019.
- [34] S. Kumar, G. Stecher, M. Li, C. Knyaz, and K. Tamura, "MEGA X: molecular evolutionary genetics analysis across computing platforms," *Molecular Biology and Evolution*, vol. 35, no. 6, pp. 1547–1549, 2018.
- [35] S.-H. Liu, C. Edwards, P. C. Hoch, P. H. Raven, and J. C. Barber, "Complete plastome sequence of *Ludwigia octovalvis*

- (Onagraceae), a globally distributed wetland plant,” *Genome Announcements*, vol. 17, no. 4, article e01274, 2016.
- [36] Q. J. Zhou, Y. M. Chen, W. Wu, R. C. Zhou, and Y. Zhang, “The complete chloroplast genome sequence of an endangered mangrove tree *Lumnitzera littorea* (Combretaceae),” *Conservation Genetics Resources*, vol. 10, no. 4, pp. 911–913, 2018.
- [37] Z. Xiao-Ming, W. Junrui, F. Li et al., “Inferring the evolutionary mechanism of the chloroplast genome size by comparing whole-chloroplast genome sequences in seed plants,” *Scientific Reports*, vol. 7, no. 1, article 1555, 2017.
- [38] H.-J. Su, S. A. Hogenhout, A. M. Al-Sadi, and C.-H. Kuo, “Complete chloroplast genome sequence of omani lime (*Citrus aurantiifolia*) and comparative analysis within the rosids,” *PLoS One*, vol. 9, no. 11, article e113049, 2014.
- [39] R. S. Millen, R. G. Olmstead, K. L. Adams et al., “Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus,” *Plant Cell*, vol. 13, no. 3, pp. 645–658, 2001.
- [40] C. Xu, W. Dong, W. Li et al., “Comparative analysis of six *Lagerstroemia* complete chloroplast genomes,” *Frontiers in Plant Science*, vol. 8, p. 15, 2017.
- [41] O. Khakhlova and R. Bock, “Elimination of deleterious mutations in plastid genomes by gene conversion,” *The Plant Journal*, vol. 46, no. 1, pp. 85–94, 2006.
- [42] W. Q. Kong and J. H. Yang, “The complete chloroplast genome sequence of *Morus cathayana* and *Morus multicaulis*, and comparative analysis within genus *Morus* L.,” *PeerJ*, vol. 8, no. 5, article e3037, 2017.
- [43] Y. Zhang, L. Li, T. L. Yan, and Q. Liu, “Complete chloroplast genome sequences of *Praxelis* (*Eupatorium catarium* Veldkamp), an important invasive species,” *Gene*, vol. 549, no. 1, pp. 58–69, 2014.
- [44] J. Saina, Z.-Z. Li, A. Gichira, and Y.-Y. Liao, “The complete chloroplast genome sequence of tree of heaven *Ailanthus altissima* (mill.) (Sapindales: Simaroubaceae), an important pantropical tree,” *International Journal of Molecular Sciences*, vol. 19, no. 4, p. 929, 2018.
- [45] J. Shaw, E. B. Lickey, E. E. Schilling, and R. L. Small, “Comparison of whole chloroplast genome sequences to choose non-coding regions for phylogenetic studies in angiosperms: the tortoise and the hare III,” *American Journal of Botany*, vol. 94, no. 3, pp. 275–288, 2007.
- [46] Y. K. Kim, S. Jo, S. H. Cheon et al., “Extensive losses of photosynthesis genes in the plastome of a mycoheterotrophic orchid, *Cyrtosia septentrionalis* (Vanilloideae: Orchidaceae),” *Genome Biology and Evolution*, vol. 11, no. 2, pp. 565–571, 2019.
- [47] B. Stoebe, W. Martin, and K. V. Kowallik, “Distribution and nomenclature of protein-coding genes in 12 sequenced chloroplast genomes,” *Plant Molecular Biology Reporter*, vol. 16, pp. 243–255, 1998.
- [48] H. Zhang, C. Li, H. Miao, and S. Xiong, “Insights from the complete chloroplast genome into the evolution of *Sesamum indicum* L.,” *PLoS One*, vol. 8, no. 11, article e80508, 2013.
- [49] S. Kikuchi, J. Bédard, M. Hirano et al., “Uncovering the protein translocon at the chloroplast inner envelope membrane,” *Science*, vol. 339, no. 6119, pp. 571–574, 2013.
- [50] W. L. Dong, R. N. Wang, N. Y. Zhang, W. B. Fan, M. F. Fang, and Z. H. Li, “Molecular evolution of chloroplast genomes of orchid species: insights into phylogenetic relationship and adaptive evolution,” *International Journal of Molecular Sciences*, vol. 19, no. 3, p. 716, 2018.
- [51] C. P. Lin, C. S. Wu, Y. Y. Huang, and S. M. Chaw, “The complete chloroplast genome of *Ginkgo biloba* reveals the mechanism of inverted repeat contraction,” *Genome Biology and Evolution*, vol. 4, no. 3, pp. 374–381, 2012.
- [52] J. Sobanski, P. Giavalisco, A. Fischer et al., “Chloroplast competition is controlled by lipid biosynthesis in evening primroses,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 12, pp. 5665–5674, 2019.
- [53] S. Wicke, G. M. Schneeweiss, C. W. dePamphilis, K. F. Müller, and D. Quandt, “The evolution of the plastid chromosome in land plants: gene content, gene order, gene function,” *Plant Molecular Biology*, vol. 76, no. 3–5, pp. 273–297, 2011.
- [54] R. Zoschke, M. Nakamura, K. Liere, M. Sugiura, T. Borner, and C. Schmitz-Linneweber, “An organellar maturase associates with multiple group II introns,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 7, pp. 3245–3250, 2010.
- [55] G. Mohr, P. S. Perlman, and A. M. Lambowitz, “Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function,” *Nucleic Acids Research*, vol. 21, no. 22, pp. 4991–4997, 1993.
- [56] X. Zhang, T. Zhou, J. Yang et al., “Comparative analyses of chloroplast genomes of Cucurbitaceae species: lights into selective pressures and phylogenetic relationships,” *Molecules*, vol. 23, no. 9, p. 2165, 2018.
- [57] Y. Qu, J. Legen, J. Arndt et al., “Ectopic transplastomic expression of a synthetic *matK* gene leads to cotyledon-specific leaf variegation,” *Frontiers in Plant Science*, vol. 9, p. 1453, 2018.
- [58] R. A. Ferreira de Lima, A. A. de Oliveira, G. D. Colletta et al., “Can plant DNA barcoding be implemented in species-rich tropical regions? A perspective from São Paulo State, Brazil,” *Genetics and Molecular Biology*, vol. 41, no. 3, pp. 661–670, 2018.
- [59] M. Rogalski, S. Ruf, and R. Bock, “Tobacco plastid ribosomal protein S18 is essential for cell survival,” *Nucleic Acids Research*, vol. 34, no. 16, pp. 4537–4545, 2006.
- [60] C. N. Fu, H. T. Li, R. Milne et al., “Comparative analyses of plastid genomes from fourteen *Cornales* species: inferences for phylogenetic relationships and genome evolution,” *BMC Genomics*, vol. 18, no. 1, pp. 956–956, 2017.
- [61] V. Kode, E. A. Mudd, S. Iamtham, and A. Day, “The tobacco plastid *accD* gene is essential and is required for leaf development,” *The Plant Journal*, vol. 44, no. 2, pp. 237–244, 2005.
- [62] Y. Madoka, K. I. Tomizawa, J. Mizoi, I. Nishida, Y. Nagano, and Y. Sasaki, “Chloroplast transformation with modified *accD* operon increases acetyl-CoA carboxylase and causes extension of leaf longevity and increase in seed yield in tobacco,” *Plant & Cell Physiology*, vol. 43, no. 12, pp. 1518–1525, 2002.
- [63] C. Chen, Y. Zheng, S. Liu et al., “The complete chloroplast genome of *Cinnamomum camphora* and its comparison with related *Lauraceae* species,” *PeerJ*, vol. 5, article e3820, 2017.
- [64] B. A. Berger, R. Kriebel, D. Spalink, and K. J. Sytsma, “Divergence times, historical biogeography, and shifts in speciation rates of Myrtales,” *Molecular Phylogenetics and Evolution*, vol. 95, pp. 116–136, 2016.