Complete chloroplast genome sequences of five *Bruguiera* species (Rhizophoraceae): comparative analysis and phylogenetic relationships

Panthita Ruang-areerate[1], Wasitthee Kongkachana[1], Chaiwat Naktang[1], Chutima Sonthirod[1], Nattapol Narong[1], Nukoon Jomchai[1], Pasin Maprasop[2], Chatree Maknual[2], Nawin Phormsin[2], Jeremy R. Shearman[1], Wirulda Pootakham[1] and Sithichoke Tangphatsornruang[1]

[1] National Omics Center, National Science and Technology Development Agency (NSTDA), Pathum Thani, Thailand
[2] Department of Marine and Coastal Resources, Bangkok, Thailand

Corresponding authors
Wirulda Pootakham,
wirulda@alumni.stanford.edu
Sithichoke Tangphatsornruang,
sithichoke.tan@nstda.or.th

ABSTRACT

Bruguiera is a genus of true mangroves that are mostly distributed in the Indo-West Pacific region. However, the number of published whole chloroplast genome sequences of *Bruguiera* species are limited. Here, the complete chloroplast sequences of five *Bruguiera* species were sequenced and assembled using Illumina data. The chloroplast genomes of *B. gymnorhiza*, *B. hainesii*, *B. cylindrica*, *B. parviflora* and *B. sexangula* were assembled into 161,195, 164,295, 164,297, 163,228 and 164,170 bp, respectively. All chloroplast genomes contain 37 tRNA and eight rRNA genes, with either 84 or 85 protein-coding genes. A comparative analysis of these genomes revealed high similarity in gene structure, gene order and boundary position of the LSC, SSC and two IR regions. Interestingly, *B. gymnorhiza* lost a *rpl32* gene in the SSC region. In addition, a *ndhF* gene in *B. parviflora* straddles both the SSC and IRB boundary regions. These genes reveal differences in chloroplast evolution among *Bruguiera* species. Repeats and SSRs in the chloroplast genome sequences were found to be highly conserved between *B. cylindrica* and *B. hainesii* as well as *B. gymnorhiza* and *B. sexangula* indicating close genetic relationships based on maternal inheritance. Notably, *B. hainesii*, which is considered a hybrid between *B. gymnorhiza* and *B. cylindrica*, appears to have inherited the chloroplast from *B. cylindrica*. Investigating the effects of selection events on shared protein-coding genes showed a positive selection in *rps7* and *rpl36* genes in all species compared to land-plant species. A phylogenetic analysis, based on 59 conserved chloroplast protein-coding genes, showed strong support that all *Bruguiera* species are in the clade Rhizophoraceae. This study provides valuable genetic information for the study of evolutionary relationships and population genetics in *Bruguiera* and other mangrove species.

INTRODUCTION

Mangroves are plants that live in coastal regions across tropical and sub-tropical regions. They play an important role in ecology by providing habitats and nurseries for many marine organisms and protecting coastlines from erosion. There are approximately 70 mangrove species belonging to 16 families (*Kathiresan & Bingham, 2001*). Rhizophoraceae is a true mangrove family consisting of ~18 species in four genera, which are *Bruguiera*, *Ceriops*, *Kandelia* and *Rhizophora* (*Tomlinson, Primack & Bunt, 1979*).

Bruguiera is a dominant and economically important mangrove genus. It produces wood that is used for fuel and house construction, and produces several useful metabolites such as diterpenes and triterpenes (*Dahdouh-Guebas et al., 2000*; *Han et al., 2004*; *Nebula, Harisankar & Chandramohanakumar, 2013*). This genus comprises six species, namely *B. cylindrica*, *B. exaristata*, *B. gymnorhiza*, *B. hainesii*, *B. parviflora* and *B. sexangula* as well as their hybrids such as *B. x rhynchopetala* and *B. x dungarra* (*Duke & Ge, 2011*; *Duke & Kudo, 2018*). Among them, *B. gymnorhiza* is the most widespread mangrove plant, while *B. hainesii* is a critically endangered mangrove plant (*FAO, 2007*; *Duke & Ge, 2011*). Most *Bruguiera* species are distributed in the Indo-West Pacific region, except *B. exaristata* that is mostly distributed in Northern Australia and Southern New Guinea (*FAO, 2007*; *Duke & Ge, 2011*; *Urashi et al., 2013*). Based on morphological characters, the genus *Bruguiera* is separated into two major groups (*Hou, 1958*; *Sheue, Yong & Yang, 2005*; *Duke & Ge, 2011*). One group (*B. gymnorhiza*, *B. exaristata* and *B. sexangula*) has large leaves and large solitary-flowered inflorescences, while the other group (*B. cylindrica*, *B. hainesii* and *B. parviflora*) has small leaves, small multiple-flowered inflorescences. In mangrove forests, the distributions of *Bruguiera* species overlap each other and it can be difficult to identify species based on morphological characters (*Abeysinghe et al., 1999*; *Duke & Ge, 2011*). Consequently, molecular markers such as random amplified polymorphic DNA (RAPD) and simple sequence repeats (SSRs) based on nuclear and chloroplast sequences have been developed for species identification in *Bruguiera* (*Abeysinghe et al., 1999*; *Islam et al., 2008*).

Chloroplasts are semi-autonomous organelles in algae, cyanobacteria and plants that are responsible for photosynthesis. Chloroplast genomes have been used to study plant species identification, population genetics, genome evolution and phylogenetics (*Shi et al., 2005*; *Dong et al., 2012*; *Saddhe, Jamdade & Kumar, 2016*; *Bi et al., 2018*; *Wu et al., 2019*). For example, DNA barcoding based on chloroplast markers such as *matK*, *rbcL* and some intergenic spacer regions (*i.e.*, *trnL-trnF* and *trnH-psbA*) is an effective tool to identify mangrove species and their evolutionary relationships (*Shi et al., 2005*; *Saddhe, Jamdade & Kumar, 2016*; *Wu et al., 2019*). In most land plants, chloroplast genomes are inherited maternally. The chloroplast structure is usually a small circular quadripartite DNA consisting of one large single copy (LSC) region, one small single copy (SSC) region and two inverted repeats (IRs), which separate the LSC and SSC (*Mower & Vickrey, 2018*). Some plant lineages have lost one copy of the IR, such as papilionoid legumes (*Palmer & Stein, 1986*; *Choi, Jansen & Ruhlman, 2019*). Over the past decade, next generation sequencing technologies have been used to generate hundreds of complete

chloroplast genomes (*Tangphatsornruang et al., 2010*; *Tangphatsornruang et al., 2011*; *Uthaipaisanwong et al., 2012*). Recently, a number of complete chloroplast genomes of mangroves in the family Rhizophoraceae were reported (*Li et al., 2019*; *Yang et al., 2019*; *Zhang, Zhong & Yuan, 2019*; *Shi et al., 2020*; *Li et al., 2021*). These mangrove chloroplast genomes are 160–164 kb in length and consist of all four regions with LSC; ~91–93 kb, SSC; ~16–19 kb and IRs; ~26–27 kb. They contain 128–131 genes and have a highly conserved organization and structure. To date, only the chloroplast genomes of *B. sexangula* and *B. gymnorhiza* are available (*Shi et al., 2020*; *Li et al., 2021*). Thus, the chloroplast genome sequences in the genus *Bruguiera* are limited.

In this study, the chloroplast genomes of five *Bruguiera* species (*B. cylindrica*, *B. gymnorhiza*, *B. hainesii*, *B. parviflora* and *B. sexangula*) in the family Rhizophoraceae were sequenced, assembled and compared to reveal their evolutionary relationship. These chloroplast genomes will provide valuable genetic information and better understanding of evolutionary relationships among *Bruguiera* species.

MATERIALS & METHODS

DNA extraction and sequencing

Fresh leaves in five *Bruguiera* species (*B. cylindrica*, *B. gymnorhiza*, *B. hainesii*, *B. parviflora*, and *B. sexangula*) were collected from Ranong province in Thailand (Table S1) and stored in liquid nitrogen. DNA samples were extracted using a standard CTAB (Cetyl Trimethyl Ammonium Bromide) method (*Doyle & Doyle, 1987*). Then the genomic DNA was sequenced using an Illumina Hiseq2000 platform according to Illumina protocols.

Chloroplast genome assembly and annotation

Based on a reference genome-based strategy, the Illumina paired-end reads and *B. sexangula* (the accession CNA0003536 in the project-ID CNP0000567 in CNSA (https://db.cngb.org/cnsa) (*Shi et al., 2020*)) were used for assembling the chloroplast genomes of the five *Bruguiera* species using GetOrganelle (*Jin et al., 2020*).

The chloroplast genome sequences were annotated using GeSeq (*Tillich et al., 2017*). The start-stop loci and intron-exon junctions of coding genes were checked manually based on reported chloroplast genes of other land-plant species. In addition to the coding genes, transfer RNAs (tRNAs) were identified using ARAGORN v1.2.36 (*Laslett & Canback, 2004*) in the GeSeq software. The circular structures of the chloroplast genomes were visualized using OGDRAW v1.3.1 (*Greiner, Lehwark & Bock, 2019*). The sequences of the five chloroplast genomes were deposited in GenBank (accession number MW836110–MW836114).

In addition, raw reads in each species were mapped to the *rpl32* sequence using BWA (*Li & Durbin, 2009*) and one primer pair was designed to amplify the *ndhF-rpl32* region (Table S2) (*Shaw et al., 2007*) in order to examine evidence for the *rpl32* gene loss in the assembled chloroplast-genome sequence of *B. gymnorhiza*.

Comparative genome analysis

The comparison of whole chloroplast genomes among the *Bruguiera* species and two close non-mangrove species, *Pellacalyx yunnanensis* (MN106253) of Rhizophoraceae and *Erythroxylum novogranatense* (NC_0306061) of Erythroxylaceae, were analyzed to investigate sequence divergence using mVISTA with the Shuffe-LAGAN mode (*Frazer et al., 2004*). The previously reported chloroplast genome of *B. sexangula* was used as a reference for comparison (*Shi et al., 2020*). Furthermore, the junctions and borders of the IR regions were illustrated using IRscope (*Amiryousefi, Hyvönen & Poczai, 2018*).

Repeat and SSR identification

Repeat sequences with length ≥50 bp were identified using REPuter (*Kurtz et al., 2001*). Furthermore, SSRs in the chloroplast genome sequences of the five *Bruguiera* species and three previously reported mangrove species, *Kandelia obavata* (MH277332), *Rhizophora x lamarckii* (NC_046517) and *Rhizophora stylosa* (MK070169), in the family Rhizophoraceae were identified using MISA (*Thiel et al., 2003*). Minimum repeat number requirements were set to 10 for mono-, five for di-, four for tri-, and three for tetra-, penta- and hexa-nucleotide repeats (*Shi et al., 2020*). The distance for compound SSRs was ≤100 bp (default).

Four polymorphic SSR loci were selected to examine polymorphism in different *Bruguiera* species. SSR primer pairs were designed for the four SSR loci and the resulting PCR product sizes were determined for each species (Table S3).

Phylogenetic analysis

To validate the phylogenetic relationships of the *Bruguiera* species based on plastid genes, 9 representative families (Rhizophoraceae, Erythroxylaceae, Ctenolophonaceae, Violaceae, Salicaceae, Passifloraceae, Euphorbiaceae, Malpighiaceae and Clusiaceae) were selected from the order Malpighiales to construct phylogenetic tree. A total of 59 conserved chloroplast coding-genes in 40 species including the five species in this study, three other mangrove species, 31 land plant species and one outgroup species in the order Oxalidales were retrieved from NCBI and used to construct a phylogenetic tree (Table S4). These genes are *atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *atpI*, *ccsA*, *cemA*, *clpP*, *matK*, *ndhA*, *ndhD*, *ndhE*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*, *petA*, *petD*, *petG*, *petL*, *petN*, *psaA*, *psaB*, *psaC*, *psaI*, *psaJ*, *psbA*, *psbC*, *psbD*, *psbF*, *psbH*, *psbJ*, *psbL*, *psbM*, *psbN*, *psbT*, *rbcL*, *rpl2*, *rpl14*, *rpl16*, *rpl33*, *rpl36*, *rpoA*, *rpoB*, *rpoC1*, *rpoC2*, *rps2*, *rps3*, *rps4*, *rps8*, *rps11*, *rps12*, *rps14*, *rps15*, *rps18*, *rps19* and *ycf3* (Table S4). Each coding-gene sequence was aligned using MUSCLE with default parameters in MEGA X (*Kumar et al., 2018*), and the aligned sequences in all genes were concatenated for each species. The best-fit model of plastid DNA substitution (GTR + I + G) was identified using the find best DNA/protein model tool in MEGA X. Based on the aligned sequences, a maximum likelihood (ML) phylogenetic tree was preformed using RAxML version 8.2.10 (*Stamatakis, 2014*) with GTRGAMMAI (GTR + I + G) model and 1,000 bootstrap values. *Averrhoa carambola* (NC_033350) in Oxalidales was used to be an outgroup species in the phylogenetic tree (*Wurdack & Davis,*
_2009_). The phylogenetic trees were visualized using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). In addition, chloroplast _rpl32_ gene loss in _B. gymnorhiza_ and other plant species in Malpighiales and Oxalidales (_Alqahtani & Jansen, 2021_) were plotted onto the their clades of the phylogenetic tree.

Gene selective pressure analysis of _Bruguiera_ chloroplasts

A total 74 shared plastid protein-coding genes in the five _Bruguiera_ species and four terrestrial plant species, _Erythroxylum novogranatense_ (NC_030601), _Ctenolophon englerianus_ (NC_049158), _Averrhoa carambola_ (NC_033350) and _Vitis rotundifolia_ (NC_023790) (Table S5), were used to investigate selection pressures. Pairwise sequence alignments for each gene between the _Bruguiera_ species and the terrestrial plant species based on the background of different evolutionary clades, including Erythroxylaceae (_E. novogranatense)_, Ctenolophonaceae (_C. englerianus_), Oxalidales (_A. carambola_), and Rosids (_Vitis rotundifolia_) (_Han et al., 2020_) were generated using MUSCLE with default setting in MEGA X (_Edgar, 2004_; _Kumar et al., 2018_). Gaps between compared sequences were removed before further analysis. Nonsynonymous (Ka) and synonymous (Ks) substitution rates (Ka/Ks) of each gene were calculated using KaKs Calculator version 2.0 (_Wang et al., 2010_), with its model-averaging method.

RESULTS

Chloroplast genome features

Five _Bruguiera_ species (_B. gymnorhiza_, _B. hainesii, B. cylindrica, B. parviflora_ and _B. sexangula_) were sequenced to generate 743,609,903–800,931,061 raw reads per species (150 bp in length) using Illumina Hiseq2000 data (Table 1). A total of 4,866,585–7,635,727 pair-end reads for each species mapped to the reference chloroplast genome of _B. sexangula_ (_Shi et al., 2020_) and were used to assemble their chloroplast genome sequences. The read coverage was estimated at over 300× for each species. The size of the assembled chloroplast genomes ranged from 161,195 (_B. gymnorhiza_) to 164,297 (_B. cylindrica_) bp (Table 1; Fig. 1). Notably, the difference in genome length of _B. cylindrica_ (164,297 bp) and _B. hainesii_ (164,295 bp) was just 2 bp. All chloroplast genome sequences formed the typical quadripartite circular structure with the four regions including one large single copy (LSC), one small single copy (SSC) and two inverted repeat regions (IRA and IRB). The length of the LSC (91,154–91,738 bp) and IR (26,299–26,393 bp) regions was highly similar in all species, whereas the size of the SSC region had the highest variation and ranged from 17,053 (_B. gymnorhiza_) to 19,956 bp (_B. sexangula_) (Table 1). The average GC content in all chloroplast genomes was approximately 35%. The GC content in the IR regions (~42%) was higher than in the LSC (~33%) and SSC (~29%) regions.

A total of 129 (_B. gymnorhiza_)–130 (_B. hainesii, B. cylindrica, B. parviflora_ and _B. sexangula_) genes including 84–85 protein-coding genes, 37 transfer (tRNA) genes and eight ribosomal RNA (rRNA) genes were identified (Fig. 1, Tables 1 and 2). Of these, 19 genes (_ndhB, rpl2, rpl23, rps7, rps12, rps19, ycf1, ycf2, rrn4.5, rrn5, rrn16, rrn23,_

Table 1 Summary of the chloroplast genome of five *Bruguiera* species.

Species	B. gymnorhiza	B. cylindrica	B. hainesii	B. parviflora	B. sexangula
Raw reads	743,609,903	791,571,475	755,597,881	781,228,900	800,931,061
Reads for chloroplast analysis	7,077,428	5,257,482	4,866,585	7,148,202	7,635,727
Base-coverage (x)	348	372	366	377	364
Genome size (bp)	161,195	164,297	164,295	163,228	164,170
LSC (bp)	91,404	91,738	91,736	91,154	91,428
SSC (bp)	17,053	19,839	19,839	19,476	19,956
IR (bp)	26,369	26,360	26,360	26,299	26,393
LSC GC content (%)	32.77	32.69	32.69	32.86	32.77
SSC GC content (%)	29.85	28.38	28.38	28.96	28.31
IR GC content (%)	42.19	42.20	42.20	42.26	42.19
Genome GC content (%)	35.54	35.22	35.22	35.43	35.25
No. of genes	129	130	130	130	130
No. of protein coding genes	82	83	83	83	83
No. of rRNAs	8	8	8	8	8
No. of tRNAs	37	37	37	37	37
No. of duplicated genes	17	17	17	17	17
Pseudogenes	2	2	2	2	2
Gene loss	*rpl32*	–	–	–	–

trnA-UGC, trnI-GAU, trnL-CAA, trnM-CAU, trnN-GUU, trnR-ACG, and *trnV-GAC*) are duplicated in the IR regions (Fig. 1 and Table 2). All photosynthesis genes, including photosystems I (five genes) and II (15 genes), NADH dehydrogenase (11 genes), cytochrome b/f complex (six genes), ATP synthase (six genes) and rubisco large subunit (*rbcL*), were found in all *Bruguiera* chloroplast genomes. Several large and small subunits of ribosomal proteins (*rpl* and *rps*) were also identified (Table 2). Interestingly, a *rpl32* gene was lost only in *B. gymnorhiza* (Tables 1 and 2). The loss was detected from low read coverage in this region of *B. gymnorhiza* (Table S6) and confirmed by lack of a PCR product from the *ndhF-rpl32* region in *B. gymnorhiza* compared with other *Bruguiera* species (Fig. S1). Furthermore, other metabolic genes including maturase (*matK*), acetyl-CoA-carboxylate (*accD*), chloroplast envelope membrane protein (*cemA*), ATP-dependent protease (*clpP*), cytochrome c synthesis (*ccsA*) and conserved open reading frames (four genes) were found in all chloroplast genomes. There are two pseudogenes (*ycf1* and *rps19*) in the chloroplast genomes (Table 2). Among all identified genes, eight protein-coding genes (*atpF, ndhA, ndhB, petB, petD, rpl2, rpl16* and *rpoC1*) contain two exons and two protein-coding genes (*rps12* and *ycf3*) contain three exons (Tables 2 and S7). The introns in these genes in the *Bruguiera* species are highly conserved (Table S7). Several tRNA genes also contain introns such as *trnA-UGC, trnC-ACA, trnI-GAU, trnK-UUU, trnL-UAA* and *trnT-CGU*. The *trnK-UUU* intron in these chloroplast genomes encodes the protein-coding gene of *matK* (Fig. 1).
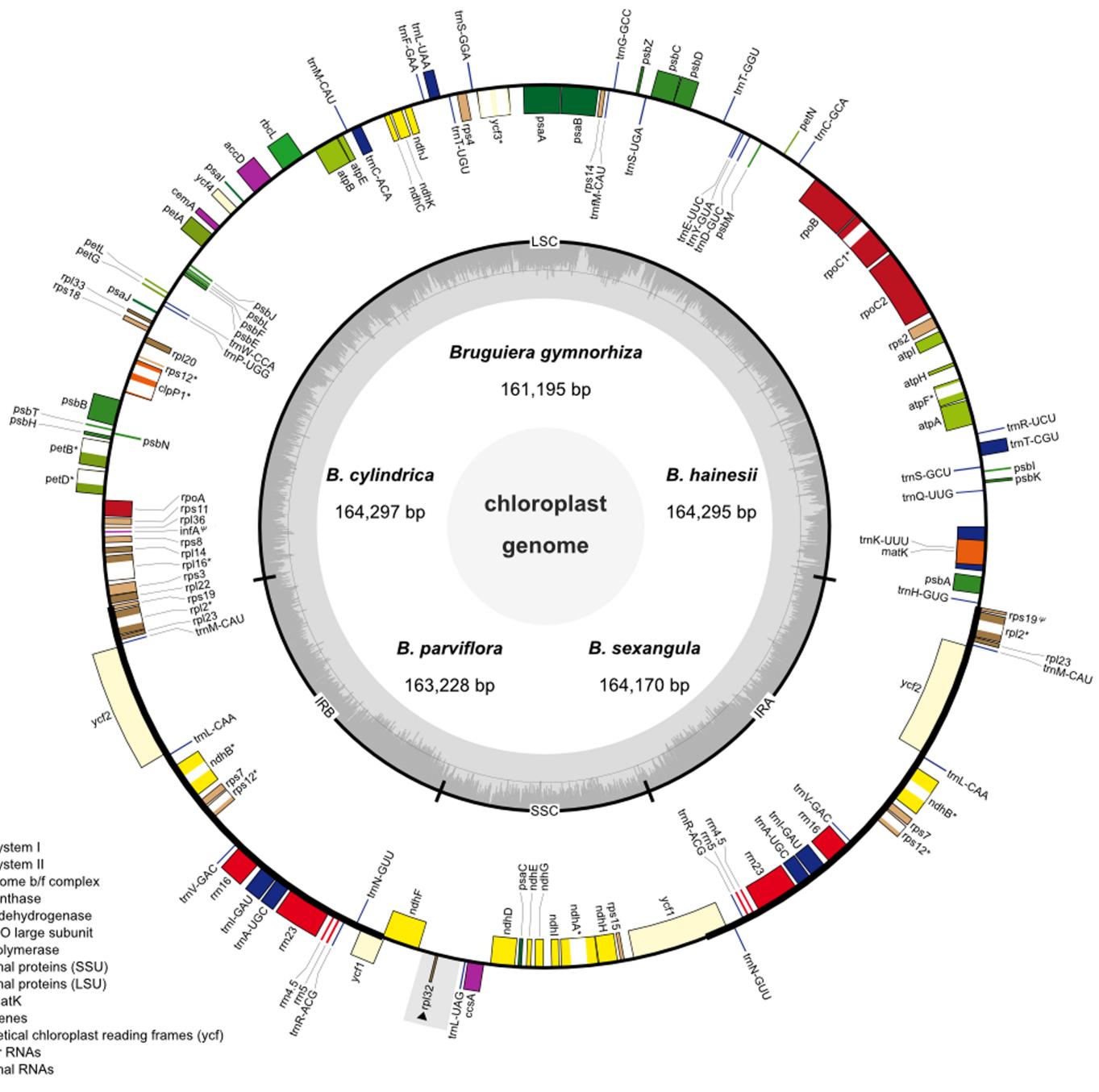
Figure 1 Circular map of the five *Bruguiera* chloroplast genomes. The inner circle shows the GC content of the chloroplast genome. LSC, SSC, and IRs (IRA and IRB) are large single copy, small single copy, and inverted repeats, respectively. Genes based on different functional groups are shown in different colors. Grey rectangle indicates a loss region of *B. gymnorhiza*.

Chloroplast genome comparison

Comparison of chloroplast genomes in the five *Bruguiera* species and two closely related non-mangrove species (*P. yunnanensis* and *E. novogranatense*) revealed similar genome organization and gene order with some sequence variation (Fig. 2). Gene orientation was

Table 2 List of annotated genes in the chloroplast genome of five *Bruguiera* species.

Category	Group of genes	Gene name
Photosynthesis	Subunits of Photosystem I	*psaA, B, C, I, J*
	Subunits of Photosystem II	*psbA, B, C, D, E, F, H, I, J, K, L, M, N, T, Z*
	Subunits of NADH dehydrogenes	*ndhA*, B*(×2), C, D, E, F, G, H, I, J, K*
	Cytochrom b6/f complex	*petA, B*, D*, G, L, N*
	ATP synthase	*atpA, B, E, F*, H, I*
	Rubisco	*rbcL*
Self-replication	Large subunit of ribosomal proteins	*rpl2*(×2), 14, 16*, 20, 22, 23(×2), 32[a], 33, 36*
	Small subunit of ribosomal proteins	*rps2, 3, 4, 7(×2), 8, 11, 12**(×2), 14, 15, 18, 19*
	DNA dependent RNA polymerase	*rpoA, B, C1*, C2*
	rRNA genes	*rrn4.5(×2), 5(×2), 16(×2), 23*(×2)*
	tRNA genes	*trnA-UGC*(×2), trnC-ACA*, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnfM-CAU, trnG-GCC, trnH-GUG,* trnI-GAU*(×2), *trnK-UUU*, trnL-CAA(×2), trnL-UAA*, trnL-UAG, trnM-CAU(×3), trnN-GUU(×2), trnP-UGG, trnQ-UUG, trnR-ACG(×2), trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-CGU*, trnT-GGU trnT-UGU, trnV-GAC(×2), trnW-CCA, trnY-GUA*
Other genes	Maturase	*matK*
	Subunit Acetyl-CoA-Carboxylate	*accD*
	Envelop membrane protein	*cemA*
	Proteaese	*clpP***
	C-type cytochrome synthesis gene	*ccsA*
Unknows	Conserved Open reading frames	*ycf1, 2(×2), 3**, 4*
	pseudogene	*ycf1, rps19*

Notes:
[*] Gene with one intron.
[**] Gene with two introns.
[a] Loss gene in *B. gymnorhiza*.

assessed by synteny analysis among the species, revealing conserved gene structure in the chloroplast genomes. The coding regions were more conserved than the intergenic regions. Both the coding and intergenic regions were highly similar in all *Bruguiera* species compared to the two non-mangrove species. The LSC and SSC regions had higher divergence than the IR regions indicating that the IR regions are highly conserved. A comparison between the two *B. sexangula* chloroplast genomes, previously reported (*Shi et al., 2020*) and the one generated in this study (Fig. 2), showed several small variations, including the difference of mononucleotide repeats, short repeats and indels. Furthermore, several poorly conserved regions located in the LSC (such as ~40.2–41.2 kb and ~55.5–56.8 kb) and SSC (such as ~120.0–121.3 kb) regions were found between the *Bruguiera* species and *E. novogranatense*. Some intergenic regions contained more variation among the species such as regions between *trnK-trnQ, trnT-trnR, trnR-atpA, trnT-psbD, psbZ-trnG, ndhC-trnC* and *ndhF-rpl32*.
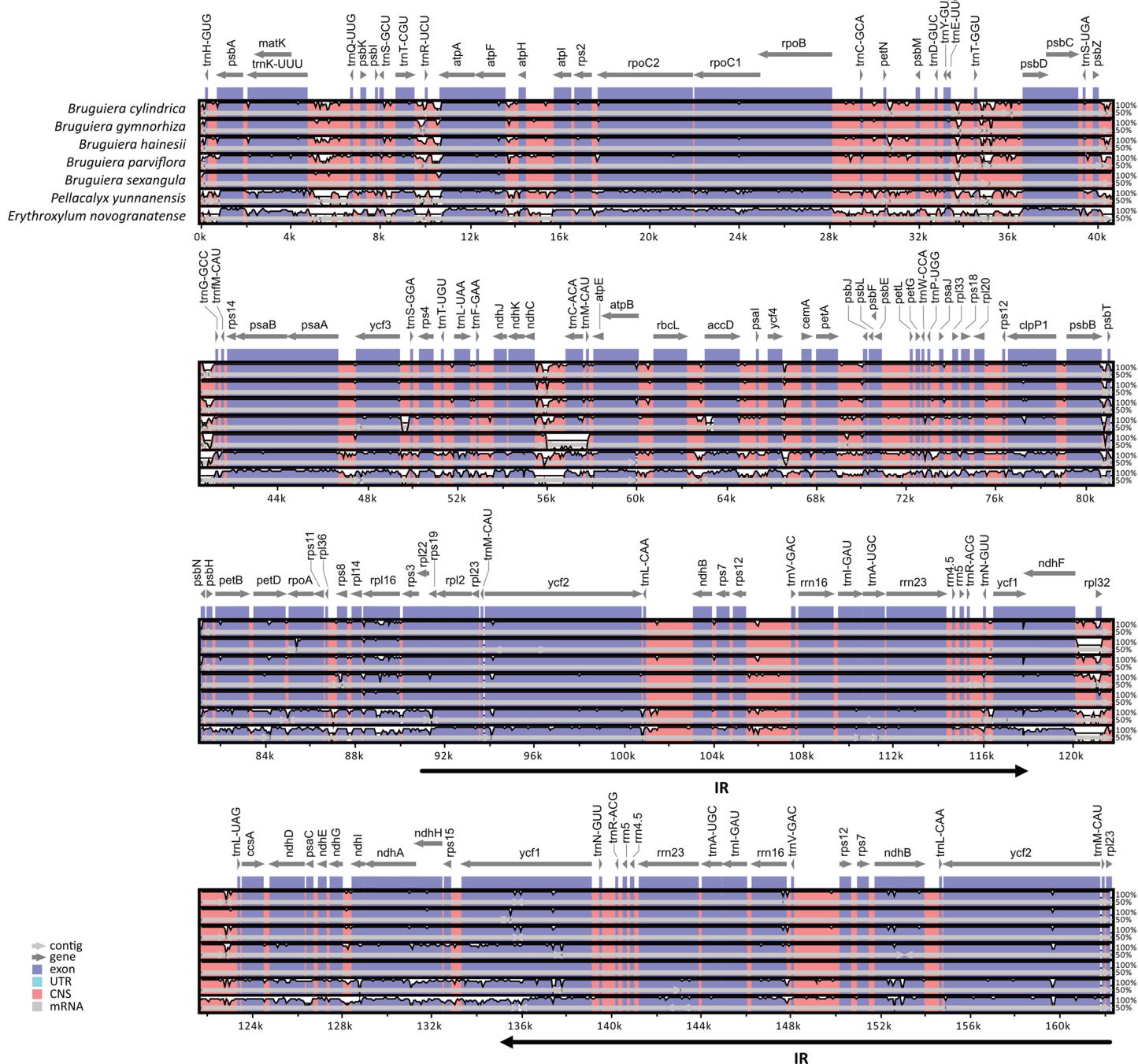
Figure 2 Alignment of the chloroplast genomes of five *Bruguiera* and two close non-mangrove species. The sequence identity was plotted among the chloroplast genomes, with *B. sexangula* (CNA0003536) as a reference using the mVista software. Grey and black arrows indicate genes and IR regions, respectively. Blue sections represent coding regions, and pink sections represent conserved non-coding sequences.

Moreover, the chloroplast boundary structures of the LSC, SSC and IRs were compared among the five species (Fig. 3). In all species, the *ycf1* gene and the *ycf1* pseudogene were located at the boundary of SSC/IRA and SSC/IRB, respectively, while the *rps19* gene and
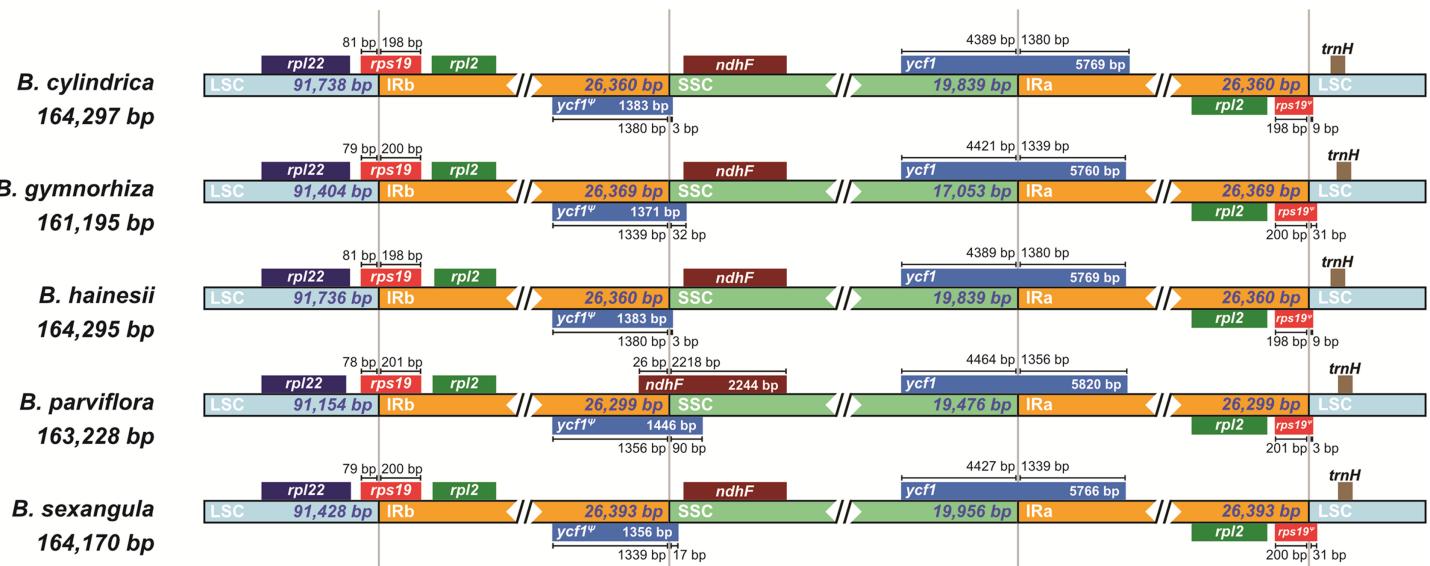
Figure 3 Comparisons of the borders of LSC, SSC, and IR regions among the five *Bruguiera* chloroplast genomes.

the *rps19* pseudogene were located at the boundary of LSC/IRB and LSC/IRA, respectively. The length of the *ycf1* and *rps19* pseudogenes varied from 1,371 bp (*B. gymnorhiza*) to 1,446 bp (*B. parviflora*) and from 204 bp (*B. parviflora*) to 231 bp (*B. gymnorhiza* and *B. sexangula*), respectively. The LSC/IR and SSC/IR borders of *B. cylindrica* and *B. hainesii* were more similar than the LSC/IR and SSC/IR borders of the other species. For example, the *ycf1* pseudogene is 3 bp away from the SSC/IRB border in both *B. cylindrica* and *B. hainesii*, whereas it is 32, 90 and 17 bp in *B. gymnorhiza*, *B. parviflora* and *B. sexangula*, respectively. The *rps19* pseudogene is 9 bp away from the LSC/IRA border in both *B. cylindrica* and *B. hainesii*, while it is 31, 3 and 31 bp in *B. gymnorhiza*, *B. parviflora* and *B. sexangula*, respectively. Additionally, the *ndhF* gene of *B. parviflora* straddles both the IRB and SSC boundary regions, with 26 bp of the IRB region and 2,218 bp of the SSC region, whereas the gene is located within the SSC region in all other species. While the sequence here is essentially the same in all species, *B. parviflora* has a variant that removes a stop codon causing *ycf1*$^\psi$ to overlap *ndhF* in the SSC/IRB boundary region.

Repeat and SSR contents

A total of 49 repeats, consisting of forward, reverse, palindromic and complement repeats, were identified in the chloroplast genome of each *Bruguiera* species (Figs. 4A–4C and Table S8). Among them, the most common repeat type was forward repeat (Fig. 4A). The repeat patterns from *B. cylindrica* and *B. hainesii* were same. Most repeats ranged in length from 24 to 30 bp (Fig. 4B) and were in the LSC region (Fig. 4C).

SSR loci in the chloroplast genomes of the five *Bruguiera* species and three previously reported mangrove species in the family Rhizophoraceae were identified (Figs. 4D–4F; Tables 3 and S9). The number of SSR loci (simple, compound, and interrupted SSRs) varied among the species, ranging from 164 in *Kandelia obavata* to 197 in *B. hainesii*
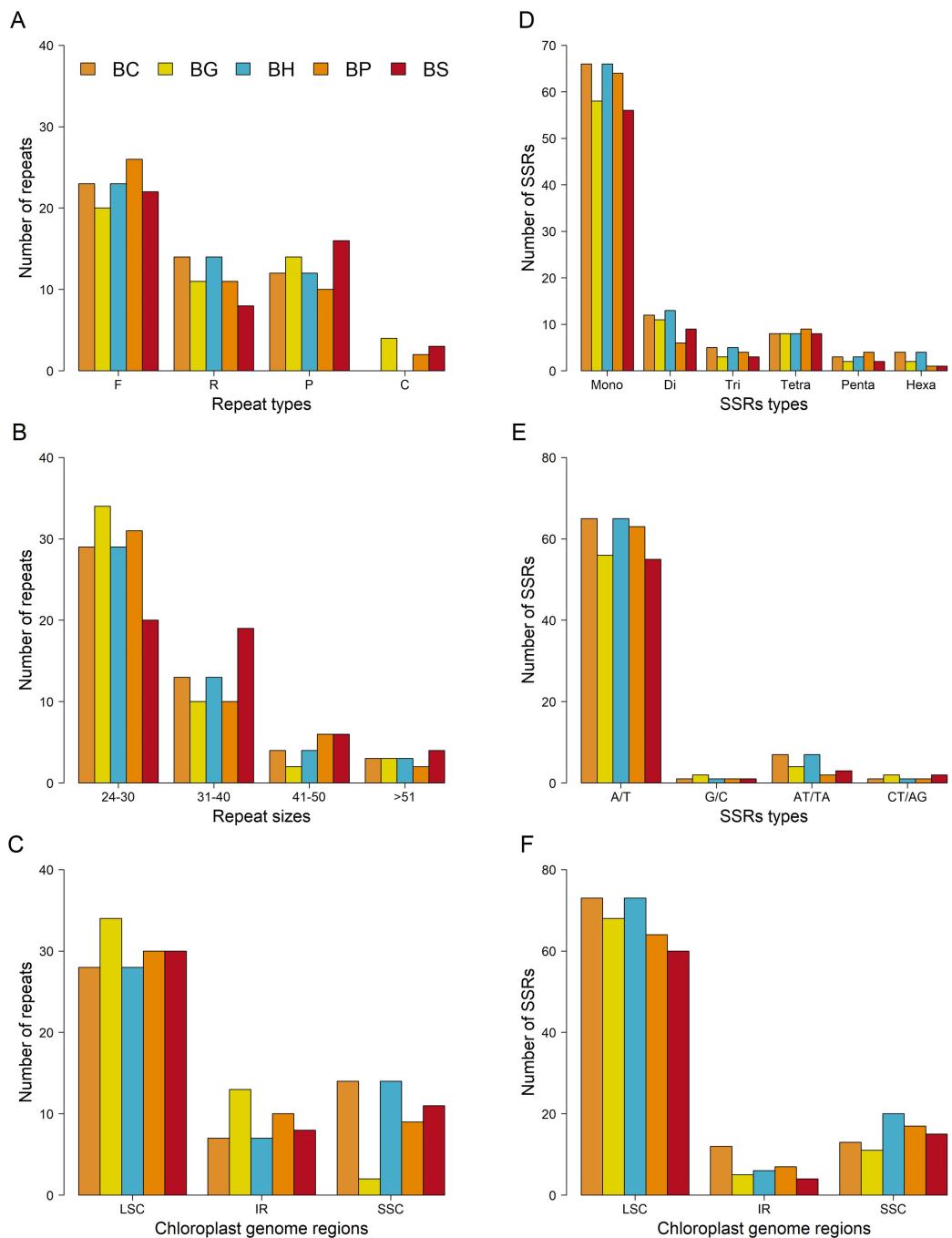
Figure 4 Comparison of repeats and microsatellites (SSRs) among the five *Bruguiera* chloroplast genomes. (A) Number of repeats based on types; (B) Number of repeats based on sizes; (C) Distribution of repeats in chloroplast regions; (D) Number of SSRs based on types; (E) Number of SSRs based on mononucleotides and dinucleotides; (F) Distribution of SSRs in chloroplast regions. BC, BG, BH, BP and BS represent *B. cylindrica*, *B. gymnorhiza*, *B. hainesii*; *B. parviflora* and *B. sexangula*, respectively.

Table 3 Number of SSRs in the chloroplast genomes of *Bruguiera* species and other reported mangrove species in the family Rhizophoraceae.

Species	SSR type						Total number	The number of SSRs for compound formation
	Mono-	Di-	Tri-	Tetra-	Penta-	Hexa-		
BH	119	31	16	16	9	6	197	61
BC	119	30	16	16	9	6	196	61
BG	102	26	14	18	4	2	166	51
BP	109	31	17	23	8	1	189	64
BS	109	36	16	19	5	1	186	70
BS[a]	106	31	14	18	5	1	175	58
KO	92	29	17	18	8	0	164	50
RL	117	22	13	10	9	2	173	52
RS	117	22	13	10	9	2	173	52

Notes:
BC, *B. cylindrica* (MW836110); BG, *B. gymnorhiza* (MW836111); BH, *B. hainesii* (MW836112); BP, *B. parviflora* (MW836113); BS, *B. sexangula* (MW836114); BS[a], *B. sexangula* (CNA0003536); KO, *Kandelia obavata* (MH277332); RL, *Rhizophora x lamarckii* (NC_046517); RS, *Rhizophora stylosa* (MK070169) and PY, *Pellacalyx yunnanensis* (MN106253).

(Table 3). The number of SSR in *B. hainesii* (197) was higher than *B. cylindrica* (186) because of a variant that increased the repeat count of an SSR to reach the minimum threshold for inclusion. Among simple SSR, the number of mononucleotide repeats accounted for ~56–68% of all loci, of which A/T repeats were the most frequent type (Figs. 4D–4E). Most of the SSR loci were in the LSC region (Fig. 4F). In addition to simple SSR loci, there were numerous compound and interrupted SSR loci (more than two different repeat types present within ≤100 bases) (Tables 3, S9 and S10). Most *Bruguiera* species have more compound and interrupted SSR loci than other mangrove species in the family Rhizophoraceae. Each SSR loci were present in many *Bruguiera* species, with the most sharing between *B. cylindrica* and *B. hainesii*, and between *B. gymnorhiza* and *B. sexangula* (Table S10; Fig. S2). Two compound SSRs with 64 and 98 bp located in noncoding and intron-*ndhA* regions, respectively, were found in all *Bruguiera* chloroplast genomes. Notably, SSR loci of the *Bruguiera* chloroplast genomes were unique compared with the other mangrove species (*K. obavata*, *R. stylosa* and *R. x lamarckii*) in the family Rhizophoraceae.

Phylogenetic relationships

Maximum likelihood (ML) phylogenetic analysis was performed based on 59 highly conserved chloroplast coding genes in 40 species including 39 plant species in Malpighiales and one outgroup species in Oxalidales (Fig. 5). The genetically distinct clades were classified into the lineages of nine families in Malpighiales consisting of nine Rhizophoraceae members, one Erythroxylaceae member, one Ctenolophonaceae member, two Violaceae members, six Salicaceae members, four Passifloraceae members, 11 Euphorbiaceae members, four Malpighiaceae members and one Clusiaceae member. The phylogenetic tree shows that all *Bruguiera* species were in the lineage Malpighiales and grouped under the clade Rhizophoraceae with a 100% bootstrap support. *B. gymnorhiza* was placed closely with *B. sexangula*, while *B. cylindrica* and *B. hainesii* were grouped
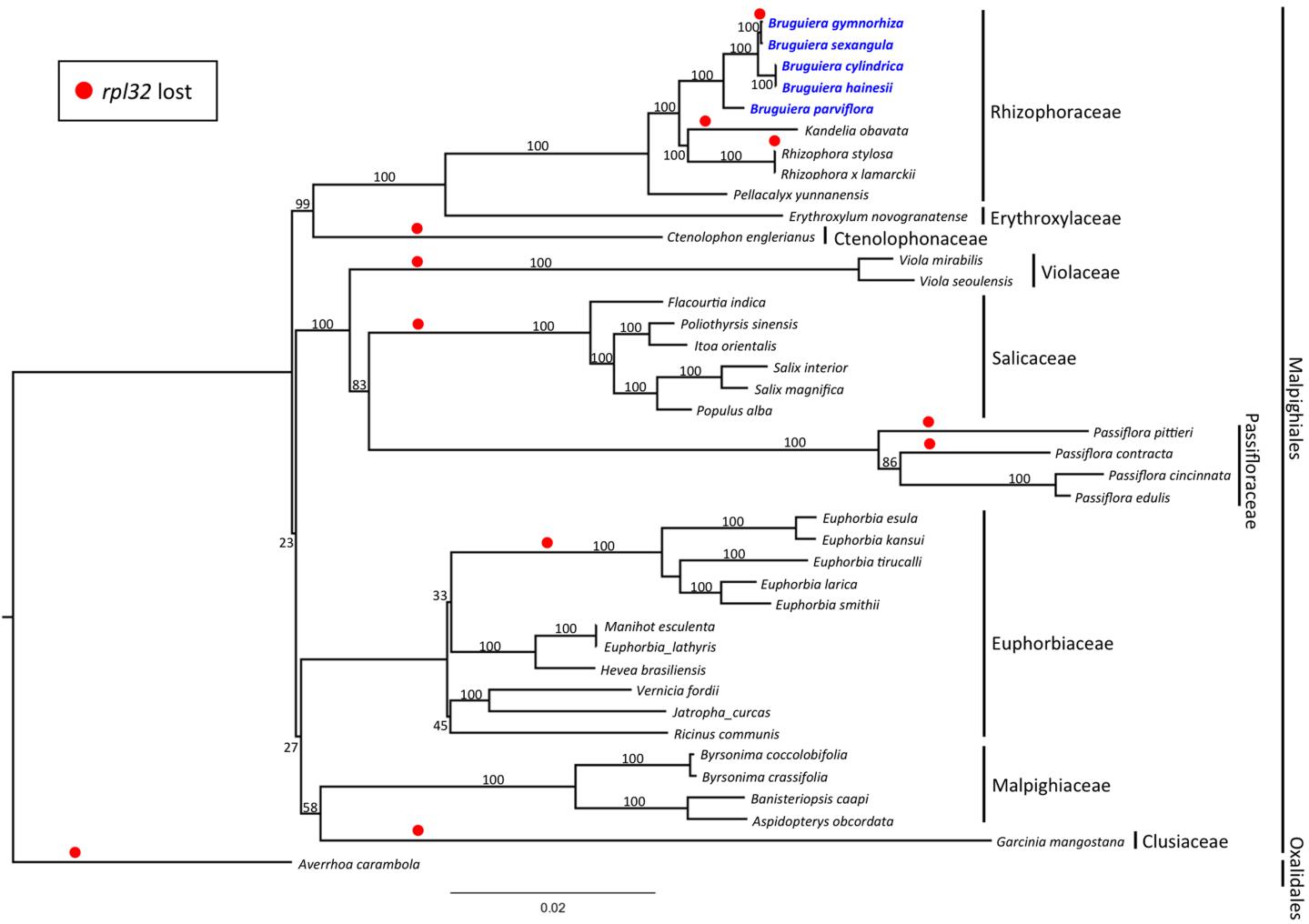
Figure 5 Molecular phylogenetic tree of the five *Bruguiera* species, three other mangrove species and thirty-two terrestrial plant species based on 59 conserved chloroplast coding genes. The maximum likelihood tree was constructed using RAxML program with the GTR + G + I nucleotide model. Numbers above nodes are bootstrap values with 1,000 replicates. *Averrhoa carambola* (NC_033350) was set as the outgroups. Loss of the *rpl32* gene is indicated by red circle.

together. *B. parviflora* is a sister species of the other four *Bruguiera* species. Loss of the *rpl32* gene was observed in several land-plant species (Fig. 5), such as *Ctenolophon engierianus*, *Viola* species, species in Salicaceae and *Euphorbia* species, and mangrove species such as *B. gymnorhiza*, *K. obavata* and *R. stylosa*.

Selective pressure genes in *Bruguiera* mangrove species

To investigate gene selection pressures, the nonsynonymous and synonymous substitution (Ka/Ks) ratio was calculated by comparing 74 conserved chloroplast gene sequences of each *Bruguiera* species to assumed ancestors, which were *E. novogranatense* (Erythroxylaceae), *C. englerianus* (Ctenolophonaceae), *A. carambola* (Oxalidales) and *V. rotundifolia* (Rosids) (Table S11). Most Ka/Ks values of compared pairs were less than 1.0. There were two genes, *rps7* and *rpl36*, in which the Ka/Ks ratios were higher than one compared with *V. rotundifolia* (Rosids) and *E. novogranatense* (Erythroxylaceae),

respectively suggesting positive selection during their evolution. The Ka/Ks ratios of the *rps7* gene between *B. cylindrica*, *B. gymnorhiza*, *B. hainesii*, *B. parviflora* and *B. sexangula* compared with *V. rotundifolia* were 1.58, 1.53, 1.58, 1.69 and 1.53, respectively. The Ka/Ks ratios of the *rpl36* gene between *B. cylindrica*, *B. gymnorhiza*, *B. hainesii*, *B. parviflora* and *B. sexangula* compared with *E. novogranatense* were 1.10, 1.10, 1.10, 0.94 and 1.10, respectively. The Ka/Ks ratios for the *rps7* gene were close to 1 when compared with *E. novogranatense* (Ka/Ks average of the five *Bruguiera* species = 0.91), *C. englerianus* (0.94), and *A. carambola* (0.97) (Table S11) revealing neutral selection within the lineages. The *rps7* gene of the *Bruguiera* species compared with *V. rotundifolia* contained 13 variation sites including 11 non-synonymous and two synonymous sites (Table 4). Five of the eleven non-synonymous sites (amino acid position: 21, 45, 105, 121, and 122) were unique to the *Bruguiera* species revealing specific nucleotide mutations during their evolution. Similar to the *rps7* gene, the *rpl36* gene of the *Bruguiera* species compared with *E. novogranatense* contained eight variation sites including six non-synonymous and two synonymous sites (Table 4). Three out of the six non-synonymous sites (amino acid position: five and 24 (two variation sites for one codon change)) were specific to the *Bruguiera* species.

DISCUSSION

In the present study, we have determined the complete chloroplast genome sequence of five *Bruguiera* species using Illumina sequence data. The chloroplast genomes have highly conserved structure and gene content, as observed in most land plants (*Daniell et al., 2016*). The size of the chloroplast genomes ranged from 161–164 kb, concordant with other mangrove species (160–164 kb) in the family Rhizophoraceae (*Li et al., 2019*; *Yang et al., 2019*; *Zhang, Zhong & Yuan, 2019*; *Shi et al., 2020*; *Li et al., 2021*). The genomes consisted of four regions (LSC, SSC and two IRs) and the IR regions were highly conserved. Interestingly, *B. gymnorhiza* (161,195 bp), which is a widespread species in the Indo-West Pacific region, had a smaller chloroplast size than the other *Bruguiera* species (163,228–164,297 bp) because of a shorter SSC region compared with the others.

In all species, the chloroplast genomes contained the same number of tRNA (37) genes, rRNA (8) genes and pseudogenes (2), whereas the number of protein-coding genes (84–85) was slightly different. The *rpl32* gene in the SSC region was lost in *B. gymnorhiza*, which was validated by PCR amplification of the *ndhF-rpl32* region (Fig. S1). Loss of the *rpl32* gene was found in at least 73 species in eudicots, occurring through multiple independent loss events, such as *K. obovata* (mangrove) and *Populus* species (*Ueda et al., 2007*; *Yang et al., 2019*; *Mohanta et al., 2020*). The transfer of the *rpl32* gene from chloroplast to the nuclear genomes occurred several times across angiosperms (*Ueda et al., 2007*; *Yang et al., 2019*; *Mohanta et al., 2020*). In this study, we found that this gene was likely transferred to the nuclear genome of *B. gymnorhiza* due to low read coverage (~100×) compared with the other *Bruguiera* species (~8,000×) (Table S6). These results suggest that the *rpl32* gene was independently lost in chloroplasts of mangrove species.

The sequence and content of the IR regions were more conserved than those of the LSC and SSC regions in the *Bruguiera* species, consistent with other mangrove chloroplast

Table 4 Variation sites in the *rps7* and *rpl36* genes between five *Bruguiera* species and four terrestrial plant species based on the reference sequence of *Vitis rotundifolia* and *Erythroxylum novogranatense*, respectively.

Gene	Nucleotide position	Nucleotide base									Amino acid position	Amino acid reference	Amino acid change
		V. rotundifolia	*A. carambola*	*C. englerianus*	*E. novogranatense*	*B. cylindrica*	*B. gymnorhiza*	*B. hainesii*	*B. parviflora*	*B. sexangula*			
rps7	62	G	G	G	G	A	A	A	G	G	21	R	Q
	99	C	C	C	T	T	T	T	T	T	33	H	H
	133	C	C	C	C	G	G	G	G	G	45	R	G
	136	T	G	G	G	G	G	G	G	G	46	S	A
	139	G	A	A	A	A	A	A	A	A	47	V	M
	167	C	C	C	A	A	A	A	A	A	56	T	K
	173	C	C	C	G	G	G	G	T	G	58	P	R/L
	187	C	C	C	T	T	T	T	T	T	63	R	C
	314	C	C	C	C	T	T	T	T	T	105	S	F
	326	C	C	C	T	T	T	T	T	T	109	P	L
	363	A	A	A	A	T	T	T	T	T	121	L	F
	364	G	G	G	G	T	T	T	T	T	122	V	L
	372	T	T	T	C	C	C	C	C	C	124	A	A
rpl36	13	G	G	G	G	A	A	G	G	A	5	A	T
	19	G	G	A	A	G	G	G	G	G	7	I	V
	44	G	G	G	G	A	A	A	G	A	15	R	Q
	55	A	A	A	A	C	C	C	C	C	19	R	R
	63	A	A	A	A	A	A	A	G	A	21	G	G
	70	A	A	A	A	T	T	T	T	T	24	I	F
	72	A	A	A	A	C	C	C	C	C	24	I	F
	82	C	T	C	C	T	T	T	T	T	28	P	S

genomes (*Li et al., 2019*; *Yang et al., 2019*; *Zhang, Zhong & Yuan, 2019*; *Shi et al., 2020*). The IR regions had high GC content, concordant with other mangrove plants (*Li et al., 2019*; *Yang et al., 2019*; *Zhang, Zhong & Yuan, 2019*; *Shi et al., 2020*). In addition, a contraction of the IR-SSC boundary region was found in *B. parviflora* (163,228 bp); as a result, its chloroplast genome size was smaller than those of *B. cylindrica*, *B. hainesii* and *B. sexangula* (164,170–164,297 bp). There was an overlap between *ycf1*$^{\Psi}$ and *ndhF* in the SSC/IRB boundary region of *B. parviflora*, similar in several plant chloroplast genomes (*Munyao et al., 2020*; *Su et al., 2020*). These are common evolutionary events that result in size variation of chloroplast genomes in land plants (*Wicke et al., 2011*). Moreover, the genes adjacent to LSC/IR and SSC/IR boundaries are *rps19*, *rps19* pseudogene, *ycf1* and *ycf1* pseudogene with variation in pseudogene length, similar in many land plants (*Gichira et al., 2019*). In contrast, the boundary of LSC/IRA junction of *K. obovata*, which is a true mangrove species in the family Rhizophoraceae, was near the *rpl22* gene resulting in a shorter LSC and larger IR to give an overall larger chloroplast genome (*Yang et al., 2019*), revealing the variation of the IR junction among mangrove species during evolution.

A large number of repeats and SSRs were identified in the chloroplast genomes, that were mostly distributed in the LSC region as this contains a large amount of intergenic sequence. No differences were found between *B. cylindrica* and *B. hainesii* for the four SSR primer pairs (Fig. S2). *B. hainesii* was reported to be a hybrid between *B. gymnorhiza* and *B. cylindrica* identified using nuclear DNA markers (*CesA* and *UKN*) (*Ono et al., 2016*). The high similarity in chloroplast sequence and lack of genetic variation between *B. cylindrica* and *B. hainesii* are consistent with the chloroplast of *B. hainesii* originating from *B. cylindrica*. Comparison of SSRs, based on the same criteria of the SSR identification, in *B. sexangula* grown in two different locations (Thailand and China) revealed intraspecific sequence variation. For example, the Thai sample contained 186 SSRs, while only 175 SSRs were previously reported in the Chinese sample using the same criteria for SSR identification (*Shi et al., 2020*). The overall pattern of chloroplast SSR loci in the *Bruguiera* species were mostly unique compared with *Kandelia* and *Rhizophora* species. Thus, these SSRs would provide a useful genetic resource for future molecular marker analyses among *Bruguiera* species.

The ML phylogenetic tree was constructed based on 59 conserved chloroplast coding-genes in 40 species. The genus *Bruguiera* and other mangrove species, *K. obovata*, *R. stylosa* and *R. x lamarckii*, constituted a monophyletic branch of the clade Rhizophoraceae. Our phylogenetic result is consistent with morphological characters and previous phylogenetic results based on nuclear DNA markers and some chloroplast genes (*e.g.*, *matK* and *rbcL*) (*Schwarzbach & Ricklefs, 2000*; *Lakshmi, Parani & Parida, 2002*; *Sheue, Yong & Yang, 2005*; *Sahoo et al., 2007*; *Duke & Ge, 2011*). For example, a phylogenetic tree was previously constructed using 34 land plant species including *Bruguiera* species (*B. cylindrica*, *B. exaristata*, *B. gymnorhiza*, *B. parviflora*, and *B. sexangula*) based on the *rbcL* gene and *trnL-trnF* spacer (*Schwarzbach & Ricklefs, 2000*). The phylogenetic result of these two sequences presented *B. gymnorhiza* and *B. sexangula* in the same clade. Another phylogenetic tree was reported based on chromosome and RAPD markers in four *Bruguiera* species that clustered into two groups with

B. gymnorhiza and *B. sexangula* in one group, and *B. cylindrica* with *B. parviflora* in the other group (*Sahoo et al., 2007*). These results confirm the close relationships of *B. gymnorhiza* and *B. sexangula*.

The Ka/Ks values for all shared protein-coding genes of the *Bruguiera* species were calculated and compared with the four assumed ancestors to assess gene selection pressure. Among the protein-coding genes, nonsynonymous substitutions (Ka) were more frequent than synonymous substitutions (Ks) in *rps7* (11 non-synonymous sites of 13 variation sites) and *rpl36* (six non-synonymous sites of eight variation sites). The Ka/Ks values of the *rps7* and *rpl36* genes were above 1.0 in all species indicating that these genes were under positive selection. The result of the *rps7* gene is consistent with previous findings for many mangrove species and some land plants such as *K. obovata*, *R. stylosa*, *R. mangle* and *Ananas comosus* (pineapple) (*Redwan, Saidin & Kumar, 2015*; *Han et al., 2020*). The *rps7* gene encoding small ribosomal protein seven is involved in the regulation of chloroplast translation (*Fargo, Boynton & Gillham, 2001*). The *rpl36* gene was reported to be under positive selection in many mangrove species such as *Pemphis acidula*, and *Sonneratia ovata* (*Shi et al., 2020*). Knockout of the *rpl36* gene in tobacco results in severe morphological aberrations, poor photoautotrophic growth and low translational efficiency (*Fleischmann et al., 2011*). Therefore, the *rps7* and *rpl36* genes might be candidates for improving stress tolerance in mangrove species in dynamic and rapidly changing coastal environments.

CONCLUSIONS

This study reports the whole chloroplast genome sequences of the five *Bruguiera* mangrove species. The comparative analyses among them provide insights into the chloroplast genome evolution in the genus *Bruguiera*. Evidence from the *Bruguiera* chloroplast genomes confirms that *B. hainesii*, a hybrid species between *B. cylindrica* and *B. gymnorhiza*, has inherited the chloroplast from *B. cylindrica*. Translocation of *rpl32* from the chloroplast to the nuclear genome was demonstrated to occur in only *B. gymnorhiza*; but not in other *Bruguiera* species. The comparison of the *Bruguiera* species in the family Rhizophoraceae based on phylogenetic relationships showed that *B. gymnorhiza* was closely related with *B. sexangula*, *B. hainesii* was confirmed to have inherited its chloroplast from *B. cylindrica*, and *B. parviflora* is a sister species of the other four *Bruguiera* species. Molecular markers were developed from SSRs and demonstrated to be useful in differentiating genotypes at both interspecific and intraspecific levels. The *rps7* and *rpl36* genes were found to be under positive selection in the *Bruguiera* species, and these genes might be important for stress tolerance. Therefore, these chloroplast genomes will provide useful genetic information for further analysis in *Bruguiera* and other mangrove species.

Funding

This work was supported by the National Science and Technology Development Agency (No. P1952261), Thailand. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Science and Technology Development Agency: P1952261.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Panthita Ruang-areerate conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Wasitthee Kongkachana analyzed the data, prepared figures and/or tables, and approved the final draft.
- Chaiwat Naktang analyzed the data, prepared figures and/or tables, and approved the final draft.
- Chutima Sonthirod analyzed the data, prepared figures and/or tables, and approved the final draft.
- Nattapol Narong performed the experiments, prepared figures and/or tables, and approved the final draft.
- Nukoon Jomchai performed the experiments, prepared figures and/or tables, and approved the final draft.
- Pasin Maprasop performed the experiments, prepared figures and/or tables, resource, and approved the final draft.
- Chatree Maknual performed the experiments, prepared figures and/or tables, resource, and approved the final draft.
- Nawin Phormsin performed the experiments, prepared figures and/or tables, resource, and approved the final draft.
- Jeremy R. Shearman analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Wirulda Pootakham conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Sithichoke Tangphatsornruang conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

All raw reads are available at Sequence Read Archive (SRA): SRR14318719 to SRR14318723. These Illumina reads were used to assemble the chloroplast genomes of five Bruguiera species.

https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA724818.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.12268#supplemental-information.

REFERENCES

Abeysinghe PD, Triest L, Greef BD, Koedam N, Hettiarachi S. 1999. Genetic differentiation between *Bruguiera gymnorhiza* and *B. sexangula* in Sri Lanka. *Hydrobiologia* **413**:11–16 DOI 10.1023/A:1003899028558.

Alqahtani AA, Jansen RK. 2021. The evolutionary fate of rpl32 and rps16 losses in the *Euphorbia schimperi* (Euphorbiaceae) plastome. *Scientific Reports* **11(1)**:1–12 DOI 10.1038/s41598-021-86820-z.

Amiryousefi A, Hyvönen J, Poczai P. 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **34(17)**:3030–3031 DOI 10.1093/bioinformatics/bty220.

Bi Y, Zhang MF, Xue J, Dong R, Du YP, Zhang XH. 2018. Chloroplast genomic resources for phylogeny and DNA barcoding: a case study on *Fritillaria*. *Scientific Reports* **8(1)**:1–12 DOI 10.1038/s41598-018-19591-9.

Choi IS, Jansen R, Ruhlman T. 2019. Lost and found: return of the inverted repeat in the legume clade defined by its absence. *Genome Biology and Evolution* **11(4)**:1321–1333 DOI 10.1093/gbe/evz076.

Dahdouh-Guebas F, Mathenge C, Kairo JG, Koedam N. 2000. Utilization of mangrove wood products around Mida Creek (Kenya) amongst subsistence and commercial users. *Economic Botany* **54(4)**:513–527 DOI 10.1007/BF02866549.

Daniell H, Lin CS, Yu M, Chang WJ. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology* **17(1)**:1–29 DOI 10.1186/s13059-016-1004-2.

Dong W, Liu J, Yu J, Wang L, Zhou S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLOS ONE* **7(4)**:1–9 DOI 10.1371/journal.pone.0035071.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**:11–15.

Duke NC, Ge XJ. 2011. *Bruguiera* (Rhizophoraceae) in the Indo-West Pacific: a morphometric assessment of hybridization within single-flowered taxa. *Blumea* **56(1)**:36–48 DOI 10.3767/000651911X572968.

Duke NC, Kudo H. 2018. *Bruguiera × dungarra*, a new hybrid between mangrove species *B. exaristata* and *B. gymnorhiza* (Rhizophoraceae) recently discovered in north-east Australia. *Blumea: Journal of Plant Taxonomy and Plant Geography* **63**:279–285 DOI 10.3767/blumea.2018.63.03.03.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32(5)**:1792–1797 DOI 10.1093/nar/gkh340.

FAO. 2007. *The world's mangroves*. Rome: Food and Agricultural Organization.

Fargo DC, Boynton JE, Gillham NW. 2001. Chloroplast ribosomal protein S7 of chlamydomonas binds to chloroplast mRNA leader sequences and may be involved in translation initiation. *The Plant Cell* **13(1)**:207–218 DOI 10.1105/tpc.13.1.207.

Fleischmann TT, Scharff LB, Alkatib S, Hasdorf S, Schottler MA, Bock R. 2011. Nonessential plastid-encoded ribosomal proteins in tobacco: a developmental role for plastid translation and implications for reductive genome evolution. *The Plant Cell* **23(9)**:3137–3155 DOI 10.1105/tpc.111.088906.

Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Research* **32(Web Server)**:W273–W279 DOI 10.1093/nar/gkh458.

Gichira AW, Avoga S, Li Z, Hu G, Wang Q, Chen J. 2019. Comparative genomics of 11 complete chloroplast genomes of Senecioneae (Asteraceae) species: DNA barcodes and phylogenetics. *Botanical Studies* **60(1)**:1–17 DOI 10.1186/s40529-019-0265-y.

Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* **47(W1)**:W59–W64 DOI 10.1093/nar/gkz238.

Han L, Huang X, Sattler I, Dahse HM, Fu H, Lin W, Grabley S. 2004. New diterpenoids from the marine mangrove *Bruguiera gymnorrhiza*. *Journal of Natural Products* **67(9)**:1620–1623 DOI 10.1021/np040062t.

Han K, Shi C, Li L, Seim I, Lee SMY, Xu X, Yang H, Fan G, Liu X. 2020. Lineage-specific evolution of mangrove plastid genomes. *Plant Genome* **13(2)**:1–13 DOI 10.1002/tpg2.20019.

Hou D. 1958. Bruguiera. *Flora Malesiana Series I* **5**:457–468.

Islam MS, Lian C, Geng Q, Kameyama N, Hogetsu T. 2008. Chloroplast microsatellite markers for the mangrove tree species *Bruguiera gymnorrhiza*, *Kandelia candel*, and *Rhizophora stylosa*, and cross-amplification in other mangrove species. *Conservation Genetics* **9(4)**:989–993 DOI 10.1007/s10592-007-9424-x.

Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* **21**:1–31 DOI 10.1101/256479.

Kathiresan K, Bingham BL. 2001. Biology of mangroves and mangrove ecosystems. *Advances in Marine Biology* **40(3)**:81–251 DOI 10.1016/S0065-2881(01)40003-4.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35(6)**:1547–1549 DOI 10.1093/molbev/msy096.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* **29(22)**:4633–4642 DOI 10.1093/nar/29.22.4633.

Lakshmi M, Parani M, Parida A. 2002. Molecular phylogeny of mangroves IX: molecular marker assisted intra-specific variation and species relationships in the Indian mangrove tribe Rhizophoreae. *Aquatic Botany* **74(3)**:201–217 DOI 10.1016/S0304-3770(02)00105-5.

Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* **32(1)**:11–16 DOI 10.1093/nar/gkh152.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25(14)**:1754–1760 DOI 10.1093/bioinformatics/btp324.

Li CT, Guo P, Huang HR, Pei NC, Shi MM, Yan HF. 2019. The complete chloroplast genome of *Rhizophora stylosa* and its phylogenetic implications. *Mitochondrial DNA Part B: Resources* **4(1)**:374–375 DOI 10.1080/23802359.2018.1547167.

Li SY, Li YC, Zhang TH, Qin LL, An YD, Pang YK, Jiang GF. 2021. Characterization of the complete chloroplast genome of mangrove *Bruguiera gymnorrhiza* (L.) Lam. ex Savigny. *Mitochondrial DNA Part B: Resources* **6(7)**:2076–2078 DOI 10.1080/23802359.2021.1914220.

Mohanta TK, Mishra AK, Khan A, Hashem A, Abd_Allah EF, Al-Harrasi A. 2020. Gene loss and evolution of the plastome. *Genes* **11**:1–23 DOI 10.1101/676304.

Mower JP, Vickrey TL. 2018. *Structural diversity among plastid genomes of land plants*. Amsterdam: Elsevier Ltd.

Munyao JN, Dong X, Yang JX, Mbandi EM, Wanga VO, Oulo MA, Saina JK, Musili PM, Hu GW. 2020. Complete chloroplast genomes of *Chlorophytum comosum* and *Chlorophytum*

gallabatense: genome structures, comparative and phylogenetic analysis. *Plants* **9(3)**:1–18 DOI 10.3390/plants9030296.

Nebula M, Harisankar HS, Chandramohanakumar N. 2013. Metabolites and bioactivities of Rhizophoraceae mangroves. *Natural Products and Bioprospecting* **3(5)**:207–232 DOI 10.1007/s13659-013-0012-0.

Ono J, Yong JWH, Takayama K, Saleh MNB, Wee AKS, Asakawa T, Yllano OB, Salmo SG, Suleiman M, Tung NX, Soe KK, Meenakshisundaram SH, Watano Y, Webb EL, Kajita T. 2016. *Bruguiera hainesii*, a critically endangered mangrove species, is a hybrid between *B. cylindrica* and *B. gymnorhiza* (Rhizophoraceae). *Conservation Genetics* **17(5)**:1137–1144 DOI 10.1007/s10592-016-0849-y.

Palmer JD, Stein DB. 1986. Conservation of chloroplast genome structure among vascular plants. *Current Genetics* **10(11)**:823–833 DOI 10.1007/BF00418529.

Redwan RM, Saidin A, Kumar SV. 2015. Complete chloroplast genome sequence of MD-2 pineapple and its comparative analysis among nine other plants from the subclass Commelinidae. *BMC Plant Biology* **15(1)**:1–20 DOI 10.1186/s12870-015-0587-1.

Saddhe AA, Jamdade RA, Kumar K. 2016. Assessment of mangroves from Goa, west coast India using DNA barcode. *Springer Plus* **5(1)**:1–10 DOI 10.1186/s40064-016-3191-4.

Sahoo P, Jena S, Mohanty S, Das AB. 2007. Molecular phylogenetic relationships among four species of the mangrove tree genus *Bruguiera* (Rhizophoraceae), as revealed by chromosome and RAPD markers. *Revista de Biologia Tropical* **55(2)**:437–448 DOI 10.15517/rbt.v55i2.6023.

Schwarzbach AE, Ricklefs RE. 2000. Systematic affinities of Rhizophoraceae and Anisophylleaceae, and intergeneric relationships within Rhizophoraceae, based on chloroplast DNA, nuclear ribosomal DNA, and morphology. *American Journal of Botany* **87(4)**:547–564 DOI 10.2307/2656599.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94(3)**:275–288 DOI 10.3732/ajb.94.3.275.

Sheue CH, Yong JWH, Yang YP. 2005. The *Bruguiera* (Rhizophoraceae) species in the mangroves of Singapore, especially on the new record and the rediscovery. *Taiwania* **50**:251–260.

Shi C, Han K, Li L, Seim I, Lee SMY, Xu X, Yang H, Fan G, Liu X. 2020. Complete chloroplast genomes of 14 mangroves: phylogenetic and genomic comparative analyses. *BioMed Research International* **2020**:1–13 DOI 10.1101/787556.

Shi S, Huang Y, Zeng K, Tan F, He H, Huang J, Fu Y. 2005. Molecular phylogenetic analysis of mangroves: independent evolutionary origins of vivipary and salt secretion. *Molecular Phylogenetics and Evolution* **34(1)**:159–166 DOI 10.1016/j.ympev.2004.09.002.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30(9)**:1312–1313 DOI 10.1093/bioinformatics/btu033.

Su T, Zhang M, Shan Z, Li X, Zhou B, Wu H, Han M. 2020. Comparative survey of morphological variations and plastid genome sequencing reveals phylogenetic divergence between four endemic Ilex species. *Forests* **11(9)**:1–17 DOI 10.3390/f11090964.

Tangphatsornruang S, Sangsrakru D, Chanprasert J, Uthaipaisanwong P, Yoocha T, Jomchai N, Tragoonrung S. 2010. The chloroplast genome sequence of mungbean (*Vigna radiata*) determined by high-throughput pyrosequencing: structural organization and phylogenetic relationships. *DNA Research* **17(1)**:11–22 DOI 10.1093/dnares/dsp025.

Tangphatsornruang S, Uthaipaisanwong P, Sangsrakru D, Chanprasert J, Yoocha T, Jomchai N, Tragoonrung S. 2011. Characterization of the complete chloroplast genome of

Hevea brasiliensis reveals genome rearrangement, RNA editing sites and phylogenetic relationships. *Gene* **475(2)**:104–112 DOI 10.1016/j.gene.2011.01.002.

Thiel T, Michalek W, Varshney RK, Graner A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* **106(3)**:411–422 DOI 10.1007/s00122-002-1031-0.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. Geseq–versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* **45(W1)**:W6–W11 DOI 10.1093/nar/gkx391.

Tomlinson PB, Primack RB, Bunt JS. 1979. Preliminary observations on floral biology in mangrove Rhizophoraceae. *Biotropica* **11(4)**:256–277 DOI 10.2307/2387918.

Ueda M, Fujimoto M, Arimura S, Murata J, Tsutsumi N, Kadowaki K. 2007. Loss of the *rpl32* gene from the chloroplast genome and subsequent acquisition of a preexisting transit peptide within the nuclear gene in *Populus*. *Gene* **402(1–2)**:51–56 DOI 10.1016/j.gene.2007.07.019.

Urashi C, Teshima KM, Minobe S, Koizumi O, Inomata N. 2013. Inferences of evolutionary history of a widely distributed mangrove species, *Bruguiera gymnorrhiza*, in the Indo-West Pacific region. *Ecology and Evolution* **3(7)**:2251–2261 DOI 10.1002/ece3.624.

Uthaipaisanwong P, Chanprasert J, Shearman JR, Sangsrakru D, Yoocha T, Jomchai N, Jantasuriyarat C, Tragoonrung S, Tangphatsornruang S. 2012. Characterization of the chloroplast genome sequence of oil palm (*Elaeis guineensis* Jacq.). *Gene* **500(2)**:172–180 DOI 10.1016/j.gene.2012.03.061.

Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. 2010. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, Proteomics and Bioinformatics* **8(1)**:77–80 DOI 10.1016/S1672-0229(10)60008-3.

Wicke S, Schneeweiss GM, DePamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology* **76(3–5)**:273–297 DOI 10.1007/s11103-011-9762-4.

Wu F, Li M, Liao B, Shi X, Xu Y. 2019. DNA barcoding analysis and phylogenetic relation of mangroves in Guangdong Province, China. *Forests* **10(1)**:1–13 DOI 10.3390/f10010056.

Wurdack KJ, Davis CC. 2009. Malpighiales phylogenetics: gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. *American Jounal of Botany* **96(8)**:1551–1570 DOI 10.3732/ajb.0800207.

Yang Y, Zhang Y, Chen Y, Gul J, Zhang J, Liu Q, Chen Q. 2019. Complete chloroplast genome sequence of the mangrove species *Kandelia obovata* and comparative analyses with related species. *PeerJ* **7(4)**:1–17 DOI 10.7717/peerj.7713.

Zhang Y, Zhong JD, Yuan CC. 2019. Complete chloroplast genome of a mangrove natural hybrid, *Rhizophora × lamarckii*. *Mitochondrial DNA Part B: Resources* **4(1)**:1465–1466 DOI 10.1080/23802359.2019.1598790.