Molecular Phylogenetics and Evolution 133 (2019) 24–32

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Enlarged and highly repetitive plastome of *Lagarostrobos* and plastid phylogenomics of Podocarpaceae

Edi Sudianto[a,b,c], Chung-Shien Wu[c], Lars Leonhard[d], William F. Martin[e,*], Shu-Miaw Chaw[a,c,*]

[a] Biodiversity Program, Taiwan International Graduate Program, Academia Sinica and National Taiwan Normal University, Taipei 11529, Taiwan
[b] Department of Life Science, National Taiwan Normal University, Taipei 11677, Taiwan
[c] Biodiversity Research Center, Academia Sinica, Taipei 11529, Taiwan
[d] Botanical Garden, Heinrich-Heine-University, 40225 Düsseldorf, Germany
[e] Institute of Molecular Evolution, Heinrich-Heine-University, 40225 Düsseldorf, Germany

ARTICLE INFO

Keywords:
Podocarp
Cupressophytes
Huon pine
Plastid genome
Plastome evolution
Repetitive DNA

ABSTRACT

Podocarpaceae is the largest family in cupressophytes (conifers II), but its plastid genomes (plastomes) are poorly studied, with plastome data currently existing for only four of the 19 Podocarpaceous genera. In this study, we sequenced and assembled the complete plastomes from representatives of eight additional genera, including *Afrocarpus, Dacrydium, Lagarostrobos, Lepidothamnus, Pherosphaera, Phyllocladus, Prumnopitys,* and *Saxegothaea*. We found that *Lagarostrobos,* a monotypic genus native to Tasmania, has the largest plastome (151,496 bp) among any cupressophytes studied to date. Plastome enlargement in *Lagarostrobos* coincides with increased intergenic spacers, repeats, and duplicated genes. Among the Podocarpaceae, *Lagarostrobos* has the most rearranged plastome, but its substitution rates are modest. Plastid phylogenomic analyses based on 81 plastid genes clarify the positions of previously conflicting Podocarpaceous genera. Tree topologies firmly support the division of Podocarpaceae into two sister clades: (1) the Prumnopityoid clade and (2) the clade containing Podocarpoid, Dacrydioid, *Pherosphaera,* and *Saxegothaea.* The *Phyllocladus* is nested within the Podocarpaceae, thus familial status of the monotypic Phyllocladaceae is not supported.

1. Introduction

Plastid genomes (plastomes) of seed plants have an average size of 145 kb and contain highly conserved structure, including two large inverted repeats (IRs) and two single copy (SC) regions (Jansen and Ruhlman, 2012). In seed plants, the smallest plastome was found in a parasitic plant, *Pilostyles* (11 kb; Bellot and Renner, 2015), while the largest one resides in *Pelargonium transvaalense* (242 kb; Tonti-Filippini et al., 2017) with the IR longer than 70 kb. Plastid gene order is highly syntenic among seed plants (Jansen and Ruhlman, 2012), except for some lineages, such as Jasminae (Lee et al., 2007), *Trifolium subterraneum* (Cai et al., 2008), *Trachelium caeruleum* (Haberle et al., 2008), Geraniaceae (Guisinger et al., 2011), and cupressophytes (Wu and Chaw, 2014, 2016; Chaw et al., 2018). In Geraniaceae, the degree of plastomic inversions is positively associated with the repeat abundance and the repeats size (Weng et al., 2014). Moreover, nucleotide substitution rates and inversion frequencies are positively correlated in the plastomes of Geraniaceae (d_N only; Weng et al., 2014) and cupressophytes (both d_N and d_S; Wu and Chaw, 2016).

Generally, plastid protein-coding genes are single copy, except those in IRs (Xiong et al., 2009). Seed plant IRs vary in size from 20 to 30 kb (Palmer, 1990), and their expansion or contraction influences the number of duplicate genes as well as plastome size. Typically, 15–17 duplicate genes are located in the IR of seed plants (Zhu et al., 2016). However, extreme IR contraction or expansion has led to only one or even > 50 genes located in the IR of Pinaceae (Wakasugi et al., 1994; Lin et al., 2010; Sudianto et al., 2016) or *Pelargonium* × *hortorum* (Chumley et al., 2006), respectively. Outside the IR, plastid gene duplication has been reported in *Pinus thunbergii* (Wakasugi et al., 1994), *Trachelium caeruleum* (Haberle et al., 2008), *Euglena archaeoplastidiata* (Bennett et al., 2017), *Monsonia emarginata* (Ruhlman et al., 2017), and *Geranium phaeum* and *G. reflexum* (Park et al., 2017). Nevertheless, mechanisms underlining gene duplication in SC regions remain poorly known.

Having lost their IR (Wu et al., 2011), cupressophytes (conifers II) offer excellent resources to study the evolution of plastome size and

* Corresponding authors at: Biodiversity Research Center, Academia Sinica, Taipei 11529, Taiwan (S.-M. Chaw). Institute of Molecular Evolution, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Duesseldorf, Germany (W.F. Martin).
E-mail addresses: bill@hhu.de (W.F. Martin), smchaw@sinica.edu.tw (S.-M. Chaw).

https://doi.org/10.1016/j.ympev.2018.12.012
Received 14 August 2018; Received in revised form 9 December 2018; Accepted 10 December 2018
Available online 13 December 2018
1055-7903/

gene duplication that are irrelevant to IR expansion or contraction. It has been revealed that in cupressophytes, synonymous substitution rates are negatively correlated with plastome size (Wu and Chaw, 2016), in agreement with the mutational burden hypothesis (Lynch, 2006). In addition, the cupressophyte plastomes are highly rearranged with the lowest GC-content among gymnosperms (Chaw et al., 2018). Only a few duplicate genes were reported in the cupressophyte plastomes, including *trnQ-UUG* (Yi et al., 2013; Guo et al., 2014), *trnN-GUU* (Vieira et al., 2014, 2016; Wu and Chaw, 2016), *trnI-CAU* and partial *rpoC2* (Hsu et al., 2016), *trnD-GUC* (Vieira et al., 2016), and *rrn5* (Wu and Chaw, 2016). Most of these duplicate genes are involved in generating family-specific short IRs that are capable of mediating homologous recombination (Wu and Chaw, 2016).

Podocarpaceae is the largest cupressophyte family, including three major clades (Dacrydioid, Podocarpoid, and Prumnopityoid), 18–19 genera, and 173–187 species (Biffin et al., 2011; Cernusak et al., 2011; Christenhusz and Byng, 2016). Podocarpaceous species are mainly distributed in the Southern Hemisphere, with the species diversity hotspot in Malesia (Enright and Jaffré, 2011). Previous phylogenetic studies were inconclusive or controversial as to the positions of some Podocarpaceous genera, such as *Lagarostrobos, Lepidothamnus, Phyllocladus,* and *Saxegothaea* (Fig. 1). For example, Conran et al. (2000) and Knopft et al. (2012) placed *Saxegothaea* as the first diverging genus of Podocarpaceae, but six other studies did not recover the same result. Other studies even classified *Saxegothaea* as a subfamily of Araucariaceae (Erdtman, 1965) or as its own family, Saxegothaeceae (Gaussen, 1973; Gajardo et al., 1996; Doweld and Reveal, 1998). Moreover, *Phyllocladus* was once treated as a monogeneric family (the so-called Phyllocladaceae), separate from Podocarpaceae on the basis of morphological characteristics, chromosome numbers, and phytochemistry (Keng, 1977, 1978; Molloy, 1996; Page, 1990). Nonetheless, various molecular phylogenetic studies congruently held the genus belonging to Podocarpaceae (Conran et al., 2000; Kelch, 2002; Biffin et al., 2011). The classification of Prumnopityoid clade also varies across studies, e.g. Sinclair et al. (2002) and Biffin et al. (2011, 2012) only included *Lagarostrobos, Manoao, Parasitaxus, Halocarpus,* and *Prumnopitys* in the clade. Some authors, however, also added *Lepidothamnus* (Knopft et al., 2012) or *Phyllocladus* (Conran et al., 2000; Knopft et al., 2012) into the clade (see Fig. 1).

Plastome diversity and evolution in the Podocarpaceae have yet to be systematically studied. To date, only the plastomes of the four genera – *Dacrycarpus, Nageia, Podocarpus*, and *Retrophyllum*, which are of the Dacrydioid and Podocarpoid clades, have been reported. However, samplings of the two clades are incomprehensive, and those of the Prumnopityoid clade have never been elucidated before. These may lead us to underestimate the plastome complexity in the Podocarpaceae. To this end, we have sequenced the complete plastomes from another eight Podocarpaceous genera with four representatives from the Prumnopityoid clade. The aim of this study was to (1) better elucidate Podocarpaceous plastome diversity and evolution, and (2)

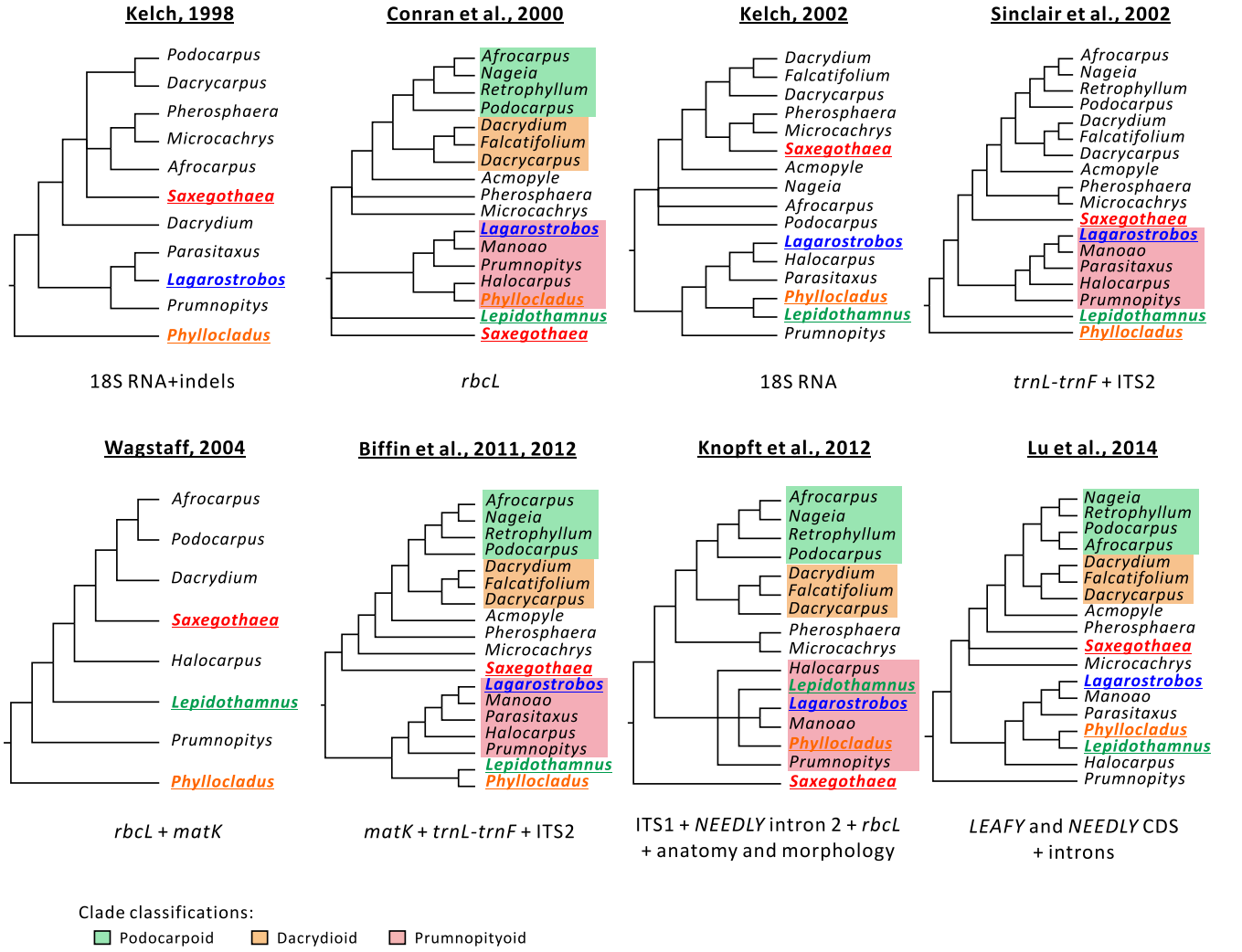

Fig. 1. Previously reported phylogeny of Podocarpaceous genera. The eight studies have competing placements of *Lagarostrobos, Lepidothamnus, Phyllocladus,* and *Saxegothaea.*

reassess intergeneric relationships across Podocarpaceae, with a particular focus on the phylogenetic position of *Phyllocladus*.

2. Materials and methods

2.1. Plant materials and DNA extraction

Plant materials were collected from eight Podocarpaceous species growing in Botanischer Garten der Heinrich-Heine-Universität Düsseldorf and University of California Botanical Garden. Table S1 summarizes their collection information, voucher numbers, and GenBank accessions. For each species, total DNA was extracted from 2 g of fresh leaves with a modified CTAB method (Stewart and Via, 1993).

2.2. Plastome sequencing, assembly, and annotation

Sequencing tasks were conducted on an Illumina NextSeq 500 platform at Genomics BioSci & Tech (New Taipei City) or Tri-I Biotech (New Taipei City) to generate approximately 2–4 Gb of 150 bp paired-end reads for each species. The software Trimmomatic 0.36 (Bolger et al., 2014) was used to remove adapters and trim the raw sequencing reads. We used Ray 2.3.1 (Boisvert et al., 2010) for *de novo* assembly. For each species, assembled scaffolds were BLAST-searched against the *Nageia nagi* plastome (AB830885), and those with *E*-value $< 10^{-10}$ were considered as plastomic scaffolds. Gaps within or between plastomic scaffolds were closed using GapFiller 1.10 (Nadalin et al., 2012) or PCR amplicons obtained from specific primers. Plastome annotations were conducted in Geneious 11.0.5 (Kearse et al., 2012) using the annotated *N. nagi* plastome as the reference, followed by manually adjusting the gene/exon boundaries. Transfer RNA (tRNA) genes were further confirmed with use of tRNAscan-SE 2.0 (Lowe and Chan, 2016). We used REPuter (Kurtz et al., 2001) to explore repetitive sequences with a minimal size of 8 bp. Pseudogenes were annotated using BLASTn search against the NCBI non-redundant (nr) database.

2.3. Sequence alignment and phylogenetic tree reconstruction

Sequences of the 81 common plastid protein-coding genes were extracted from the eight newly sequenced and other publicly available plastomes, including four Podocarpaceous and three Araucariaceous species (Table S1). Each gene was aligned using MUSCLE (Edgar, 2004) implemented in MEGA 7.0 (Kumar et al., 2016), with the "Align Codon" option. We used SequenceMatrix (Vaidya et al., 2011) to concatenate the alignments, yielding a supermatrix that contained 84,798 characters for tree construction. Maximum likelihood (ML) and Bayesian inference (BI) trees were estimated using RAxML 8.2.10 (Stamatakis, 2014) and MrBayes 3.2.6 (Ronquist and Huelsenbeck, 2003) under a GTR + G model suggested by jModelTest 2.1.10 (Darriba et al., 2012), respectively. The bootstrap support (BS) on ML tree was assessed from 1,000 replicates. The BI analysis was run for 2,000,000 generations, in which a tree was sampled per 100 generations. The first 25% of the sampled trees were discarded as the burn-in, while the remaining ones were used to calculate the Bayesian posterior probabilities (PP).

2.4. Identification of locally collinear blocks and reconstruction of ancestral plastomes

The software progressiveMauve (Darling et al., 2010) was used to identify locally collinear blocks (LCBs) between the 12 Podocarpaceous species and *Cycas taitungensis* plastomes (AP009339). IR_A sequences were removed from the *Cycas* plastome prior to the LCB identification as previous studies indicated that the cupressophyte plastomes have lost IR_A (Wu et al., 2011; Wu and Chaw, 2014). Matrix of the LCB order and ML tree were used to reconstruct the ancestral gene order by using RINGO (Feijão and Araujo, 2016). GRIMM (Tesler, 2002) was used to estimate the plastomic inversion history.

2.5. Estimation of nucleotide substitution rates

Both non-synonymous (d_N) and synonymous (d_S) substitution rates were calculated using CODEML program in PAMLX (Xu and Yang, 2013). The parameters set were runmode = 0, seqtype = 1, CodonFreq = 2, estFreq = 0, model = 1, and cleandata = 1. The ML tree shown in Fig. 3 was the constraint tree.

2.6. Analyses of divergence times and absolute substitution rates

The relative divergence times of the Podocarpaceous species were estimated using RelTime (Tamura et al., 2012) in MEGA 7.0. We used five estimated points from TimeTree (Hedges et al., 2015) as the calibrated ages of five specific nodes (Fig. S6). Absolute d_N and d_S substitution rates (R_N and R_S) were obtained by dividing d_N and d_S branch lengths by the estimated time along the corresponding branches. Only branches leading to the extant species were taken into account.

3. Results

3.1. Lagarostrobos plastome is enlarged with abundant repeats

The eight newly sequenced plastomes are 130,343–151,496 bp long, with GC content ranging from 36.6 to 37.6%. They contain 81–83 protein-coding, 33–35 tRNA, and four rRNA genes (Table 1; Fig. S1). Among them, the *Lagarostrobos* plastome is the largest. However, its gene density (0.81 genes per kb) is lower than those of other seven taxa (0.86–0.91 genes per kb). As a result, *Lagarostrobos* has the highest content of intergenic spacer (IGS) among the elucidated taxa (Table 1). *Lagarostrobos* also contains unusually abundant repeats, occupying 22.2% of its plastome (Table 1; Fig. 2). By contrast, the repeat content was estimated to be 1.1–5.3% in other Podocarpaceous plastomes, suggesting that proliferation of repeats is highly active in the *Lagarostrobos* plastome. Collectively, the increased IGS and repeat contents together contribute to the enlarged plastome in *Lagarostrobos*.

Table 1
Features of the eight newly sequenced plastomes of Podocarpaceous species.

Species	Size (bp)	GC (%)	No. of genes				Gene density (genes/kb)	Non-genic sequences (%)		Repeats (%)
			Protein-coding	tRNA	rRNA	Pseudogenes		Introns	IGS	
Afrocarpus gracilior	134,317	37.3	81	34	4	2	0.89	8.2	30.8	5.3
Dacrydium cupressinum	131,809	37.3	81	34	4	3	0.90	8.3	29.8	3.4
Lagarostrobos franklinii	151,496	36.6	83	35	4	30	0.81	7.7	36.8	22.2
Lepidothamnus intermedius	130,343	37.3	81	33	4	3	0.91	8.9	28.7	1.1
Pherosphaera fitzgeraldii	135,722	37.6	81	35	4	1	0.88	8.2	30.7	4.4
Phyllocladus aspleniifolius	136,466	37.0	81	32	4	4	0.86	8.6	30.1	4.4
Prumnopitys andina	137,117	37.2	81	34	4	3	0.87	8.6	32.0	5.1
Saxegothaea conspicua	134,130	37.5	81	34	4	4	0.89	8.8	30.5	4.1

IGS: intergenic spacers.

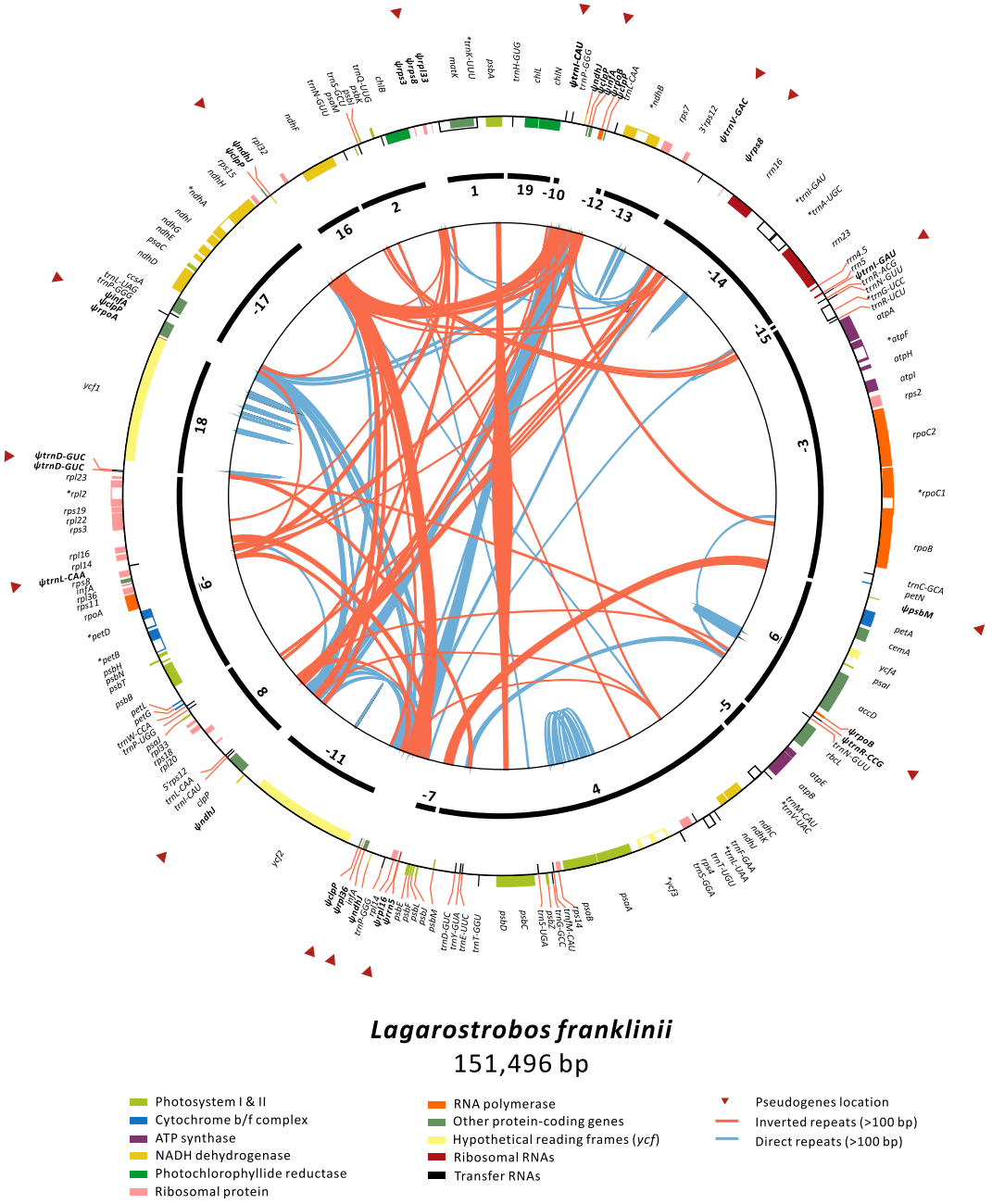

Fig. 2. Repeat- and pseudogene-rich plastome of *Lagarostrobos franklinii*. Colored boxes on the outermost circle represent genes being transcribed in clockwise (inner boxes) or counterclockwise (outer boxes) directions. Pseudogenes are labeled with a psi (Ψ). Black lines with numbers in the middle depict the locally collinear blocks (LCBs) between *Cycas* and Podocarpaceous plastomes. Negative signs indicate opposite direction as compared to *Cycas*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A number of plastid genes vary greatly in lengths across Podocarpaceae. For example, *Lagarostrobos* and *Phyllocladus* have expanded *accD* that are 1.2 and 1.7-times longer than the average of other taxa (2,033 bp), respectively (Fig. S2). In addition, the *clpP* gene is elongated in *Lagarostrobos* and *Pherosphaera* due to insertions of repetitive sequences at the 3′ end (Fig. S3). Whereas, loss of introns has led to shrinkage of the *rpoC1* sequences in *Afrocarpus*, *Dacrycarpus*, *Dacrydium*, *Nageia*, *Podocarpus*, and *Retrophyllum* (Fig. S1). This suggests that loss of *rpoC1* intron occurred before the divergence of the Podocarpoid and Dacrydioid clades. By contrast, *Pherosphaera* is the sole taxon lacking the *atpF* intron (Fig. S1).

3.2. Outburst of duplicate genes and pseudogenes in the Lagarostrobos plastome

Only *Lagarostrobos* has duplicated *rpl14* and *infA* (Fig. 2; Fig. S4). Distinct substitution rates were observed between the paralogs of these two genes, suggesting divergent evolution after the gene duplication. *Lagarostrobos* has experienced several lineage-specific duplications of plastid tRNAs, including three functional *trnP-GGG*, two functional and one pseudo (*Ψ*) *trnL-CAA*, and one functional and one *ΨtrnI-CAU* (Fig. S1). In addition, *Lagarostrobos* has two *ΨtrnD-GUC* copies that are tandemly arranged.

Overall, the pseudogene content in the *Lagarostrobos* plastome amounts to ca. 3.9 kb and includes partial sequences from 22 protein-coding genes, 7 tRNAs, and 1 rRNA (Fig. 2; Table 1; Table S2). Whereas,

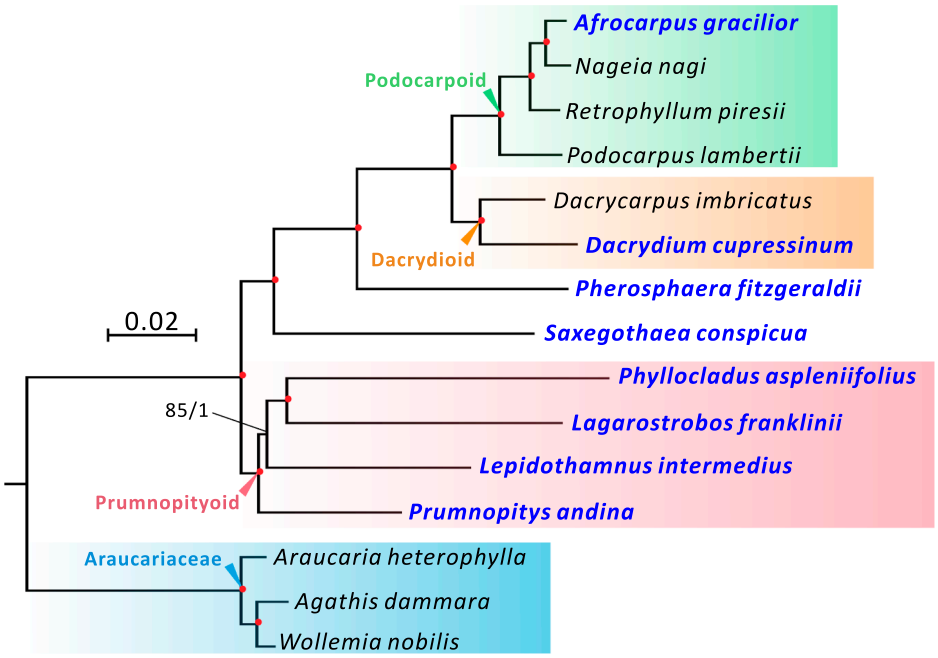

Fig. 3. Plastid phylogenomics of Podocarpaceae. The tree framework is based on the ML tree inferred from 81 plastid protein-coding genes with three Araucariaceous species as the outgroup. Numbers along nodes indicate bootstrap support (BS)/posterior probability (PP) for ML/BI analyses. Red dots signify full support for both analyses. Newly sequenced species are highlighted in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

only 1–4 pseudogenes were detected in other confamilial taxa (Table 1). Most of the pseudogenes in *Lagarostrobos* are lineage-specific; only *ΨrpoB* and *ΨtrnD-GUC* are shared with other Podocarpaceous lineages (Fig. S1). In *Lagarostrobos*, *ΨclpP* is the most copy-number abundant, with six copies scattered over the plastome (Table S2). They roughly match with three distinct regions of the functional *clpP* (Fig. S5), with the mutation rates varying from 0.061 to 0.156 mutations per site.

Taken together, the *Lagarostrobos* plastome had multiple rounds of gene duplication and pseudogenization, which are unprecedented not only within Podocarpaceae but also among cupressophyte families.

3.3. Plastid phylogenomic analysis resolved two major clades in Podocarpaceae

Fig. 3 shows ML and BI trees inferred from 81 protein-coding genes in 12 Podocarpaceous and three Araucariaceous species. These two trees are congruent in topology with 100% bootstrap support (BS) and 1.0 posterior probability (PP) in most of the nodes. Two major sister clades were resolved: (1) the Prumnopityoid clade, including *Phyllocladus, Lagarostrobos, Lepidothamnus,* and *Prumnopitys* and (2) the clade containing Podocarpoid (*Nageia, Afrocarpus, Retrophyllum,* and *Podocarpus*), Dacrydioid (*Dacrycarpus* and *Dacrydium*), *Pherosphaera,* and *Saxegothaea*. Within Prumnopityoid, *Prumnopitys* is the earliest-diverging genus, followed by *Lepidothamnus* and the clade consisting of *Lagarostrobos* and *Phyllocladus*. The sister-relationship of the latter two genera disagrees with the Keng's views that treated *Phyllocladus* as a monogeneric family (Phyllocladaceae; Keng, 1977, 1978; also see Fig. 1). *Saxegothaea* and *Pherosphaera* are the successive sister taxa to the Podocarpoid-Dacrydioid clade. Within Podocarpoid, *Podocarpus* is the earliest divergent lineage, followed by *Retrophyllum* and the *Afrocarpus-Nageia* clade.

3.4. Podocarpaceae plastomes contain extensive inversions and diverse sets of intermediate-sized repeats

Seventeen locally collinear blocks were identified between 12 Podocarpaceous species and *Cycas* (Fig. 4). The putative LCB order at each internode was inferred on the basis of the tree shown in Fig. 3. Plastomic inversions that a species has experienced were estimated by comparing its LCB order with that of its ancestors. Our results show that *Lagarostrobos* has undergone at least five plastomic inversions after it diverged from *Phyllocladus* (Fig. 4). Three inversions have occurred specifically in *Pherosphaera*, while *Lepidothamnus* and *Podocarpus* each have undergone a lineage-specific inversion. No inversion was detected between *Saxegothaea* and the putative common ancestor of Podocarpaceae, thereby suggesting that the *Saxegothaea* plastome likely has retained the ancestral gene order in Podocarpaceae.

In Podocarpaceous plastomes, we found several intermediate-sized (100–1,000 bp; Alverson et al., 2011) repeats located near the boundary of LCBs (Fig. 4). All sampled taxa share an inverted repeat (IR) and a direct repeat (DR) with duplicated *trnN-GUU* (hereafter called *trnN*-IR) and *trnD-GUC* (*trnD*-DR), respectively. The *trnN*-IR and *trnD*-DR were previously reported as recombination substrates in Podocarpaceae (Vieira et al., 2016; Wu and Chaw, 2016). Here, we identified a number of novel repeats that are shared in specific clades, including *trnD-trnY*-DR in most Podocarpaceous genera, *chlB*-DR in the Dacrydioid clade, *trnL-trnI*-IR in *Pherosphaera*, *rpl22*-IR in *Saxegothaea*, and *petD*-DR and *chlB*-IR in *Prumnopitys* (Fig. 4). Overall, our data unravel a diverse set of intermediate-sized repeats in Podocarpaceae, in which *Lagarostrobos* contains the most abundant and complex repeats that are aggregated in pseudogene-rich regions (Fig. 2).

3.5. Rearrangement frequencies do not coincide with nucleotide substitution rates in Podocarpaceae plastomes

Our dating analysis indicates that diversification of Podocarpaceae occurred around 13.39–142 MYA (Fig. S6), in agreement with the viewpoint that the crown group of Podocarpaceae diversified during Mid-Jurassic to Mid-Cretaceous (Biffin et al., 2012). The absolute non-synonymous (R_N) and synonymous (R_S) substitution rates were estimated based on 81 concatenated genes across Podocarpaceae. In Fig. 5, the estimated R_N and R_S rates vary from 0.13 to 0.36 substitutions per site per billion years (SSB) and from 0.43 to 0.92 SSB, respectively. *Phyllocladus* has the highest R_N rate (0.36 SSB) but its R_S rate (0.69 SSB) is modest, resulting in its R_N/R_S ratio (0.53) being the largest among Podocarpaceae. In contrast, *Nageia* has the fastest R_S (0.92 SSB), while *Prumnopitys* has the slowest rates at both non-synonymous (R_N = 0.13 SSB) and synonymous (R_S = 0.43 SSB) sites. Nevertheless, none of the estimated R_N/R_S ratios of the 12 examined Podocarpaceous lineages exceeds 1, even for the highly rearranged *Lagarostrobos* (red dot in

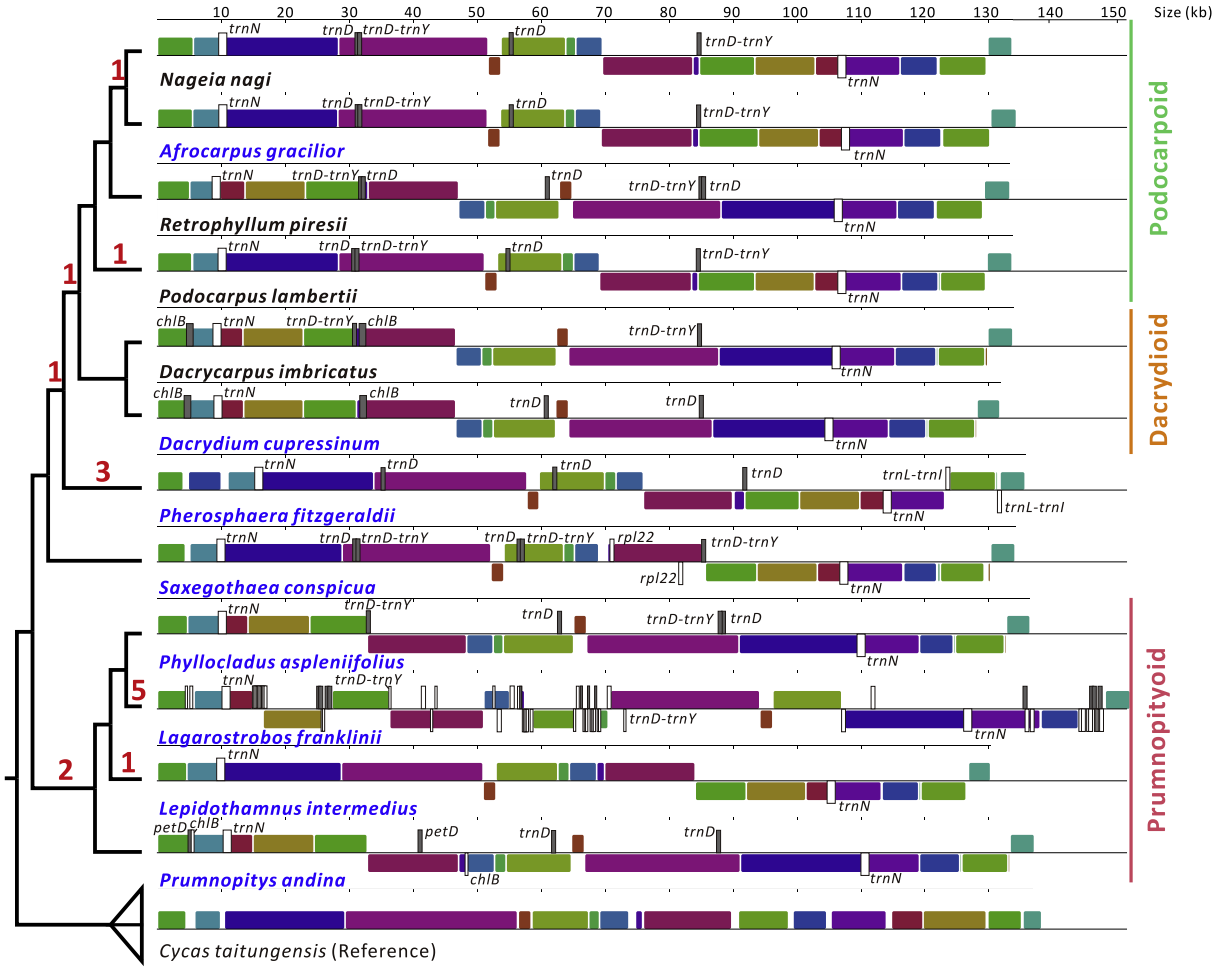

Fig. 4. Extensive plastome rearrangements during the evolution of Podocarpaceae. The tree topology on the left side was modified from Fig. 3, while colored boxes on the right side are LCBs between *Cycas* and Podocarpaceous plastomes. Boxes below horizontal lines suggest an opposing orientation relative to their counterparts in *Cycas*. Direct and inverted repeats longer than 100 bp are labeled in gray and white, respectively. Genes or gene fragments inside repeats are shown. Estimated numbers of inversion events are indicated on tree branches. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

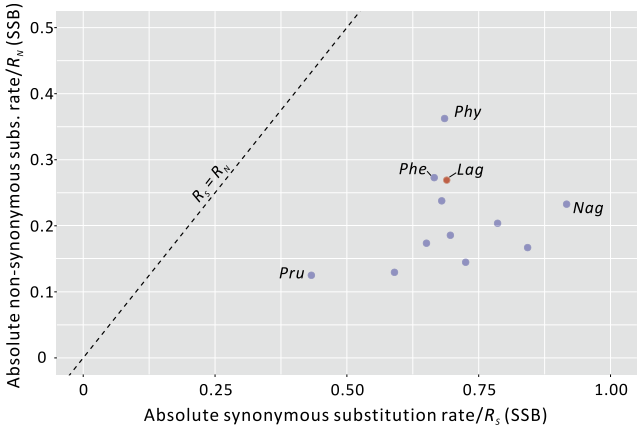

Fig. 5. Comparison of absolute non-synonymous (R_N) and synonymous substitution rates (R_S) estimated from plastid genes across the 12 Podocarpaceous species. The substitution rates of *Lagarostrobos* plastid genes are marked in red. SSB, substitutions per site per billion years; *Lag, Lagarostrobos franklinii; Phe, Pherosphaera fitzgeraldii; Nag, Nageia nagi; Phy, Phyllocladus aspleniifolius; Pru, Prumnopitys andina.* (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 5). This finding indicates that most coding sequences of Podocarpaceous plastomes are under functional constraints, irrespective of the frequency in genome rearrangements.

4. Discussion

4.1. Insights into the phylogeny of Podocarpaceae

Most molecular phylogenetic studies of Podocarpaceae in the past were based on maximum parsimony (MP) trees inferred from a few plastid loci, e.g. *rbcL, matK, trnL-trnF,* or some nuclear genes and introns (Fig. 1). But they could not agree on the intergeneric relationships in Podocarpaceae. For example, the MP trees generated by Conran et al. (2000), Sinclair et al. (2002), and Knopf et al. (2012) placed *Lepidothamnus* and/or *Saxegothaea* as the earliest divergent lineages in Podocarpaceae. This was refuted by others that placed *Phyllocladus* (Wagstaff, 2004) or *Prumnopitys* (Lu et al., 2014) as the first diverging genus. Biffin et al. (2011, 2012) divided the Podocarpaceae into two major clades on the basis of three loci (Fig. 1) but was still cautious about the positions of *Lepidothamnus* and *Phyllocladus* in the family.

We used 81 plastid genes to infer the phylogeny of Podocarpaceae – the largest dataset to date. Both our ML and BI trees strongly support a sister relationship between the Prumnopityoid clade and the clade containing Podocarpoid, Dacrydioid, *Pherosphaera,* and *Saxegothaea*

(Fig. 3), and include *Prumnopitys, Lepidothamnus, Lagarostrobos,* and *Phyllocladus* in the Prumnopityoid clade – in agreement with the classification of Knopf et al. (2012), but not those of Conran et al. (2000) and Biffin et al. (2011). Therefore, our plastid phylogenomics suggests *Phyllocladus* to be a genus of Podocarpaceae rather than constituting the monotypic Phyllocladaceae. Our tree also infers that *Saxegothaea* and *Pherosphaera* are the successive sisters to the Podocarpoid-Dacrydioid clade, contradicting the views of Conran et al. (2000) and Knopf et al. (2012) but agreeing with those of Sinclair et al. (2002) and Biffin et al. (2011, 2012) – see Fig. 1. The intergeneric relationships within the Podocarpoid and Dacrydioid clades are highly consistent with those inferred from other studies (Conran et al., 2000; Sinclair et al., 2002; Biffin et al., 2011; Knopf et al., 2012). In future analyses, including more Podocarpaceous taxa is needed to confirm the intergeneric relationships inferred herein.

4.2. Mechanisms underlying the enlarged Lagarostrobos plastome

We discovered that *Lagarostrobos* has a highly rearranged and enlarged plastome that contains abundant repeats (Table 1). Previously, the cupressophyte plastomes were documented to be 121.1–147.7 kb in size (Wu and Chaw, 2016). With our newly sequenced data, we revised the largest of the cupressophyte plastomes to be 151.5 kb with the *Lagarostrobos* plastome holding the record.

Two mechanisms have been said to account for the enlargement of plastomes: (1) expansion of IRs (Chumley et al., 2006) and (2) accumulation of IGS sequences (Smith, 2018). We ruled out the first mechanism because *Lagarostrobos* lacks IRs (Fig. 2). Instead, *Lagarostrobos* has more abundant repeats and IGS than any other Podocarpaceae (Table 1). Repeats play an important role in plastomic rearrangements (Weng et al., 2014; Sveinsson and Cronk, 2014). In addition, genomic rearrangements have been associated with increased IGS content (Sloan et al., 2012; Wu and Chaw, 2014). In *Lagarostrobos*, some long non-syntenic IGS regions (e.g. regions between LCB 11–7 and 12–10 in Fig. 2) contain rich repeats, implying that they are byproducts of repeat-mediated rearrangements. We also found that, in *Lagarostrobos,* many duplicate genes are located in the repeat-rich regions (Fig. 2), highlighting the key role of repeat proliferation in plastid gene duplication. Moreover, insertion of repetitive sequences accounts for the elongation of *accD* and *clpP* in *Lagarostrobos* (Figs. S2 and S3). As a result, repeats have led to an increase in IGS content, generation of numerous duplicate genes, and elongation of some coding genes, which together contribute to the enlarged *Lagarostrobos* plastome.

4.3. Abundant intermediate-sized repeats are likely responsible for numerous rearrangements in Lagarostrobos

Intermediate-sized repeats are able to trigger low-frequency recombination in plant mitochondrial genomes, resulting in the accumulation of substoichiometric genomes (termed sublimons), which are present at lower levels compared to the main genome (Woloszynska, 2010). The sublimons, however, may become predominant via a process known as substoichiometric shifting (SSS). Mitochondrial SSS is frequently reported in both natural and cultivated plant populations (Woloszynska, 2010). Plastomic sublimons have been reported in a number of cupressophytes, including Cupressaceae (Guo et al., 2014; Qu et al., 2017), Sciadopityaceae (Hsu et al., 2016), and Podocarpaceae (Vieira et al., 2016), and its shift is also evident among the four *Juniperus* plastomes (Guo et al., 2014). In the *Lagarostrobos* plastome, intermediate-sized repeats are far more abundant than any other Podocarpaceous genera. The abundant repeat facilitates the formation of more sublimons in *Lagarostrobos* than other genera. Notably, most of these repeats are situated at the LCB junctions, reinforcing the vital role of repeats in plastome rearrangements. Because of the presence of numerous sublimons in *Lagarostrobos,* we infer that there were at least five SSS events during the evolution of the *Lagarostrobos* plastome (Fig. 4), making it as the most rearranged plastome in Podocarpaceae.

How does *Lagarostrobos* maintain such a large plastome? Smith (2016) associates bloated organellar genomes with slow mutation rates (μ) and small effective population sizes (N_e). A recent study also indicates an inverse relationship between plastome size and d_S rate in cupressophytes (Wu and Chaw, 2016). Yet, *Lagarostrobos*' R_s is not the lowest among Podocarpaceae (Fig. 5). But as a long-lived, isolated, and highly inbred species (Shapcott, 1997), *Lagarostrobos* has an exceptionally small π (Clark and Carbone, 2008). As the substitution rates of *Lagarostrobos* plastid genes are moderate (Fig. 5), we expect the small π value was mainly due to the small N_e. A small N_e impedes the ability of natural selection to purge off excess noncoding DNA, ultimately resulting in enlarged genomes (Lynch, 2006). This viewpoint accounts for the unusually large genomes maintained in volvocine green algae (Smith et al., 2013) and similarly for the large *Lagarostrobos* plastome reported here.

Acknowledgments

We are grateful to Dr. Sabine Etges, Botanical Garden, Heinrich-Heine-University for the kind permission in collecting samples, and to Holly Forbes, Clare W. Loughran, and Bee-Yau Huang for the help in collecting the samples in UC Botanical Garden at Berkeley. This work was supported by research grants from the ERC (666053 to WFM), the Ministry of Science and Technology Taiwan (MOST 103-2621-B-001-007-MY3 to SMC) and Biodiversity Research Center's PI grant (2016–2018 to SMC), and the Taiwan International Graduate Program Student Fellowship (to ES).

Declarations of interest

None.

Data deposition

The complete plastid genomes of eight Podocarpaceous genera have been deposited at DDBJ under the accession numbers AP018899–AP018906.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2018.12.012.

Reference

Alverson, A.J., Zhuo, S., Rice, D.W., Sloan, D.B., Palmer, J.D., 2011. The mitochondrial genome of the legume *Vigna radiata* and the analysis of recombination across short mitochondrial repeats. e16404. PLoS One 6. https://doi.org/10.1371/journal.pone.0016404.

Bellot, S., Renner, S.S., 2015. The plastomes of two species in the endoparasite genus *Pilostyles* (Apodanthaceae) each retain just five or six possibly functional genes. Genome Biol. Evol. 8, 189–201. https://doi.org/10.1093/gbe/evv251.

Bennett, M.S., Shiu, S.-H., Triemer, R.E., 2017. A rare case of plastid protein-coding gene duplication in the chloroplast genome of *Euglena archaeoplastidiata* (Euglenophyta). J. Phycol. 53, 493–502. https://doi.org/10.1111/jpy.12531.

Biffin, E., Brodribb, T.J., Hill, R.S., Thomas, P., Lowe, A.J., 2012. Leaf evolution in Southern Hemisphere conifers tracks the angiosperm ecological radiation. Proc. Biol. Sci. 279, 341–348. https://doi.org/10.1098/rspb.2011.0559.

Biffin, E., Conran, J.G., Lowe, A.J., 2011. Podocarp evolution: A molecular phylogenetic perspective. Smithsonian Contributions to Botany 1–20. https://doi.org/10.5479/si.0081024X.95.1.

Boisvert, S., Laviolette, F., Corbeil, J., 2010. Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. J. Comput. Biol. 17, 1519–1533. https://doi.org/10.1089/cmb.2009.0238.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Cai, Z., Guisinger, M., Kim, H.-G., Ruck, E., Blazier, J.C., McMurtry, V., Kuehl, J.V., Boore, J., Jansen, R.K., 2008. Extensive reorganization of the plastid genome of *Trifolium subterraneum* (Fabaceae) is associated with numerous repeated sequences and novel

DNA insertions. J. Mol. Evol. 67, 696–704. https://doi.org/10.1007/s00239-008-9180-7.
Cernusak, L.A., Adie, H., Bellingham, P.J., Biffin, E., Brodribb, T.J., Coomes, D.A., Dalling, J.W., Dickie, I.A., Enright, N.J., Kitayama, K., Ladd, P.G., Lambers, H., Lawes, M.J., Lusk, C.H., Morley, R.J., Turner, B.L., 2011. Podocarpaceae in tropical forests: a synthesis. Smithsonian Contrib. Bot. 189–195. https://doi.org/10.5479/si.0081024X.95.189.
Chaw, S.-M., Wu, C.-S., Sudianto, E., 2018. Evolution of gymnosperm plastid genomes. In: Plastid Genome Evolution, Advances in Botanical Research. Elsevier, pp. 195–222. https://doi.org/10.1016/bs.abr.2017.11.018.
Christenhusz, M.J.M., Byng, J.W., 2016. The number of known plants species in the world and its annual increase. Phytotaxa 261, 201. https://doi.org/10.11646/phytotaxa.261.3.1.
Chumley, T.W., Palmer, J.D., Mower, J.P., Fourcade, H.M., Calie, P.J., Boore, J.L., Jansen, R.K., 2006. The complete chloroplast genome sequence of *Pelargonium* x *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol. Biol. Evol. 23, 2175–2190. https://doi.org/10.1093/molbev/msl089.
Clark, C.M., Carbone, I., 2008. Chloroplast DNA phylogeography in long-lived Huon pine, a Tasmanian rain forest conifer. Can. J. For. Res. 38, 1576–1589. https://doi.org/10.1139/X07-209.
Conran, J.G., Wood, G.M., Martin, P.G., Dowd, J.M., Quinn, C.J., Gadek, P.A., Price, R.A., 2000. Generic relationships within and between the gymnosperm families Podocarpaceae and Phyllocladaceae based on an analysis of the chloroplast gene *rbcL*. Aust. J. Bot. 48, 715–724.
Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. e11147. PLoS One 5. https://doi.org/10.1371/journal.pone.0011147.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109.
Doweld, A.B., Reveal, J.L., 1998. Validation of new suprageneric names in Pinophyta. Phytologia 84, 363–367.
Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinf. 5, 113. https://doi.org/10.1186/1471-2105-5-113.
Enright, N.J., Jaffré, T., 2011. Ecology and distribution of the malesian podocarps. Smithsonian Contributions to Botany 57–77. https://doi.org/10.5479/si.0081024X.95.57.
Erdtman, G., 1965. Pollen and spore morphology/plant taxonomy. Gymnospermae, Bryophyta (text). Almqvist & Wiksell, Stockholm.
Feijão, P., Araujo, E., 2016. Fast ancestral gene order reconstruction of genomes with unequal gene content. BMC Bioinf. 17, 413. https://doi.org/10.1186/s12859-016-1261-9.
Gajardo, R., Woltz, P., Gondran, M., Marguerier, J., 1996. Xylologie des conifères endémiques des Andes méridionales au MEB. I. Saxegothaeaceae. Revue de Cytologie et de Biologie Végétales – Le Botaniste 19, 31–45.
Gaussen, H., 1973. Les gymnospermes actuelles et fossiles. Fascicule XII. Les podocarpines. Étude générale. Travaux du Laboratoire Forestière de Toulouse tome 2, vol. 1, fasc. XII. Toulouse: Faculté des Sciences.
Guisinger, M.M., Kuehl, J.V., Boore, J.L., Jansen, R.K., 2011. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. Mol. Biol. Evol. 28, 583–600. https://doi.org/10.1093/molbev/msq229.
Guo, W., Grewe, F., Cobo-Clark, A., Fan, W., Duan, Z., Adams, R.P., Schwarzbach, A.E., Mower, J.P., 2014. Predominant and substoichiometric isomers of the plastid genome coexist within *Juniperus* plants and have shifted multiple times during cupressophyte evolution. Genome Biol. Evol. 6, 580–590. https://doi.org/10.1093/gbe/evu046.
Haberle, R.C., Fourcade, H.M., Boore, J.L., Jansen, R.K., 2008. Extensive rearrangements in the chloroplast genome of *Trachelium caeruleum* are associated with repeats and tRNA genes. J. Mol. Evol. 66, 350–361. https://doi.org/10.1007/s00239-008-9086-4.
Hedges, S.B., Marin, J., Suleski, M., Paymer, M., Kumar, S., 2015. Tree of life reveals clock-like speciation and diversification. Mol. Biol. Evol. 32, 835–845. https://doi.org/10.1093/molbev/msv037.
Hsu, C.-Y., Wu, C.-S., Chaw, S.-M., 2016. Birth of four chimeric plastid gene clusters in Japanese umbrella pine. Genome Biol. Evol. 8, 1776–1784. https://doi.org/10.1093/gbe/evw109.
Jansen, R.K., Ruhlman, T.A., 2012. Plastid genomes of seed plants. In: Bock, R., Knoop, V. (Eds.), Genomics of Chloroplasts and Mitochondria, Advances in Photosynthesis and Respiration. Springer, Netherlands, Dordrecht, pp. 103–126 doi:10.1007/978-94-007-2920-9_5.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
Kelch, D.G., 2002. Phylogenetic assessment of the monotypic genera *Sundacarpus* and *Manoao* (Coniferales: Podocarpaceae) utilising evidence from 18S rDNA sequences. Aust. Syst. Bot.
Knopft, P., Schulz, C., Little, D.P., Stützel, T., Stevenson, D.W., 2012. Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. Cladistics 28, 271–299. https://doi.org/10.1111/j.1096-0031.2011.00381.x.
Keng, H., 1977. *Phyllocladus* and its bearing on the systematics of conifers, In: Flowering Plants: Evolution and Classification of Higher Categories, Springer-Verlag, 235–251.
Keng, H., 1978. The genus *Phyllocladus* (Phyllocladaceae). J. Arnold Arboretum 59, 249–273.
Knopf, P., Schulz, C., Little, D.P., Stützel, T., Stevenson, D.W., 2012. Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. Cladistics 28, 271–299. https://doi.org/10.1111/j.1096-0031.2011.00381.x.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.
Kurtz, S., Choudhuri, J.V., Ohlebusch, E., Schleiermacher, C., Stoye, J., Giegerich, R., 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 29, 4633–4642.
Lee, H.-L., Jansen, R.K., Chumley, T.W., Kim, K.-J., 2007. Gene relocations within chloroplast genomes of *Jasminum* and *Menodora* (Oleaceae) are due to multiple, overlapping inversions. Mol. Biol. Evol. 24, 1161–1180. https://doi.org/10.1093/molbev/msm036.
Lin, C.-P., Huang, J.-P., Wu, C.-S., Hsu, C.-Y., Chaw, S.-M., 2010. Comparative chloroplast genomics reveals the evolution of Pinaceae genera and subfamilies. Genome Biol. Evol. 2, 504–517. https://doi.org/10.1093/gbe/evq036.
Lowe, T.M., Chan, P.P., 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. 44, W54–W57. https://doi.org/10.1093/nar/gkw413.
Lu, Y., Ran, J.-H., Guo, D.-M., Yang, Z.-Y., Wang, X.-Q., 2014. Phylogeny and divergence times of gymnosperms inferred from single-copy nuclear genes. e107679. PLoS One 9. https://doi.org/10.1371/journal.pone.0107679.
Lynch, M., 2006. Streamlining and simplification of microbial genome architecture. Annu. Rev. Microbiol. 60, 327–349. https://doi.org/10.1146/annurev.micro.60.080805.142300.
Molloy, B.P.J., 1996. A new species name in *Phyllocladus* (Phyllocladaceae) from New Zealand. New Zeal. J. Bot. 34, 287–297. https://doi.org/10.1080/0028825X.1996.10410695.
Nadalin, F., Vezzi, F., Policriti, A., 2012. GapFiller: a *de novo* assembly approach to fill the gap within paired reads. BMC Bioinf. 13 (Suppl 14), S8. https://doi.org/10.1186/1471-2105-13-S14-S8.
Page, C.N., 1990. Phyllocladaceae. In: Kramer, K.U., Green, P.S. (Eds.), Pteridophytes and Gymnosperms. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 317–319 doi:10.1007/978-3-662-02604-5_57.
Palmer, J.D., 1990. Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet. 6, 115–120. https://doi.org/10.1016/0168-9525(90)90125-P.
Park, S., Ruhlman, T.A., Weng, M.-L., Hajrah, N.H., Sabir, J.S.M., Jansen, R.K., 2017. Contrasting patterns of nucleotide substitution rates provide insight into dynamic evolution of plastid and mitochondrial genomes of *Geranium*. Genome Biol. Evol. 9, 1766–1780. https://doi.org/10.1093/gbe/evx124.
Qu, X.-J., Wu, C.-S., Chaw, S.-M., Yi, T.-S., 2017. Insights into the existence of isomeric plastomes in Cupressoideae (Cupressaceae). Genome Biol. Evol. https://doi.org/10.1093/gbe/evx071.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180.
Ruhlman, T.A., Zhang, J., Blazier, J.C., Sabir, J.S.M., Jansen, R.K., 2017. Recombination-dependent replication and gene conversion homogenize repeat sequences and diversify plastid genome structure. Am. J. Bot. 104, 559–572. https://doi.org/10.3732/ajb.1600453.
Shapcott, A., 1997. Population genetics of the long-lived Huon pine *Lagarostrobos franklinii*: An endemic Tasmanian temperate rainforest tree. Biol. Conserv. 80, 169–179. https://doi.org/10.1016/S0006-3207(96)00076-6.
Sinclair, W.T., Mill, R.R., Gardner, M.F., Woltz, P., Jaffré, T., Preston, J., Hollingsworth, M.L., Ponge, A., Möller, M., 2002. Evolutionary relationships of the New Caledonian heterotrophic conifer, *Parasitaxus usta* (Podocarpaceae), inferred from chloroplast trn L-F intron/spacer and nuclear rDNA ITS2 sequences. Plant Syst. Evol. 233, 79–104. https://doi.org/10.1007/s00606-002-0199-8.
Sloan, D.B., Alverson, A.A., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D., Taylor, D.B., 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PloS Biol. 10, e1001241. https://doi.org/10.1371/journal.pbio.1001241.
Smith, D.R., 2016. The mutational hazard hypothesis of organelle genome evolution: 10 years on. Mol. Ecol. 25, 3769–3775. https://doi.org/10.1111/mec.13742.
Smith, D.R., 2018. Plastid genomes hit the big time. New Phytol. 219, 491–495. https://doi.org/10.1111/nph.15134.
Smith, D.R., Hamaji, T., Olson, B.J.S.C., Durand, P.M., Ferris, P., Michod, R.E., Featherston, J., Nozaki, H., Keeling, P.J., 2013. Organelle genome complexity scales positively with organism size in volvocine green algae. Mol. Biol. Evol. 30, 793–797. https://doi.org/10.1093/molbev/mst002.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
Stewart, C.N., Via, L.E., 1993. A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. Biotechniques 14, 748–750.
Sudianto, E., Wu, C.-S., Lin, C.-P., Chaw, S.-M., 2016. Revisiting the plastid phylogenomics of pinaceae with two complete plastomes of *Pseudolarix* and *Tsuga*. Genome Biol. Evol. 8, 1804–1811. https://doi.org/10.1093/gbe/evw106.
Sveinsson, S., Cronk, Q., 2014. Evolutionary origin of highly repetitive plastid genomes within the clover genus (*Trifolium*). BMC Evol. Biol. 14, 228. https://doi.org/10.1186/s12862-014-0228-6.
Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., Kumar, S., 2012.

Estimating divergence times in large molecular phylogenies. Proc. Natl. Acad. Sci. U. S. A. 109, 19333–19338. https://doi.org/10.1073/pnas.1213199109.
Tesler, G., 2002. GRIMM: genome rearrangements web server. Bioinformatics 18, 492–493.
Tonti-Filippini, J., Nevill, P.G., Dixon, K., Small, I., 2017. What can we do with 1000 plastid genomes? Plant J. 90, 808–818. https://doi.org/10.1111/tpj.13491.
Vaidya, G., Lohman, D.J., Meier, R., 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171–180. https://doi.org/10.1111/j.1096-0031.2010.00329.x.
Vieira, L. do N., Faoro, H., Rogalski, M., Fraga, H.P., Cardoso, R.L., de Souza, E.M., de Oliveira Pedrosa, F., Nodari, R.O., Guerra, M.P., 2014. The complete chloroplast genome sequence of *Podocarpus lambertii*: genome structure, evolutionary aspects, gene content and SSR detection. PLoS One 9https://doi.org/10.1371/journal.pone.0090618. e90618.
Vieira, L., Rogalski, M., Faoro, H., Pacheco de Freitas Fraga, H., Goulart dos Anjos, K., Assine Picchi, G.F., Onofre Nodari, R., de Oliveira Pedrosa, F., Maltempi de Souza, E., Pedro Guerra, M., 2016. The plastome sequence of the endemic Amazonian conifer, *Retrophyllum piresii* (Silba) C.N.Page, reveals different recombination events and plastome isoforms. Tree Genet. Genomes 12 (10). https://doi.org/10.1007/s11295-016-0968-0.
Wagstaff, S.J., 2004. Evolution and biogeography of the austral genus *Phyllocladus* (Podocarpaceae). J. Biogeogr. 31, 1569–1577. https://doi.org/10.1111/j.1365-2699.2004.01066.x.
Wakasugi, T., Tsudzuki, J., Ito, S., Shibata, M., Sugiura, M., 1994. A physical map and clone bank of the black pine (*Pinus thunbergii*) chloroplast genome. Plant Mol. Biol. Rep. 12, 227–241. https://doi.org/10.1007/BF02668746.
Weng, M.-L., Blazier, J.C., Govindu, M., Jansen, R.K., 2014. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. Mol. Biol. Evol. 31, 645–659. https://doi.org/10.1093/molbev/mst257.
Woloszynska, M., 2010. Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes–though this be madness, yet there's method in't. J. Exp. Bot. 61, 657–671. https://doi.org/10.1093/jxb/erp361.
Wu, C.-S., Chaw, S.-M., 2014. Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): evolution towards shorter intergenic spacers. Plant Biotechnol. J. 12, 344–353. https://doi.org/10.1111/pbi.12141.
Wu, C.-S., Chaw, S.-M., 2016. Large-scale comparative analysis reveals the mechanisms driving plastomic compaction, reduction, and inversions in Conifers II (Cupressophytes). Genome Biol. Evol. 8, 3740–3750. https://doi.org/10.1093/gbe/evw278.
Wu, C.-S., Wang, Y.-N., Hsu, C.-Y., Lin, C.-P., Chaw, S.-M., 2011. Loss of different inverted repeat copies from the chloroplast genomes of Pinaceae and cupressophytes and influence of heterotachy on the evaluation of gymnosperm phylogeny. Genome Biol. Evol. 3, 1284–1295. https://doi.org/10.1093/gbe/evr095.
Xiong, A.-S., Peng, R.-H., Zhuang, J., Gao, F., Zhu, B., Fu, X.-Y., Xue, Y., Jin, X.-F., Tian, Y.-S., Zhao, W., Yao, Q.-H., 2009. Gene duplication, transfer, and evolution in the chloroplast genome. Biotechnol. Adv. 27, 340–347. https://doi.org/10.1016/j.biotechadv.2009.01.012.
Xu, B., Yang, Z., 2013. PAMLX: a graphical user interface for PAML. Mol. Biol. Evol. 30, 2723–2724. https://doi.org/10.1093/molbev/mst179.
Yi, X., Gao, L., Wang, B., Su, Y.-J., Wang, T., 2013. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): evolutionary comparison of *Cephalotaxus* chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms. Genome Biol. Evol. 5, 688–698. https://doi.org/10.1093/gbe/evt042.
Zhu, A., Guo, W., Gupta, S., Fan, W., Mower, J.P., 2016. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. New Phytol. 209, 1747–1756. https://doi.org/10.1111/nph.13743.