HAYATI
Journal of Biosciences

Vol. 33 No. 1, January 2026 240-249
DOI:10.4308/hjb.33.1.240-249
ISSN: 1978-3019
EISSN: 2086-4094

Short Communication

Evolutionary Conservation and Diversity of Interspecific and Intraspecific Callophylaceae Chloroplast Genomes

Nhat Nam Nguyen[1], Hoang Dang Khoa Do[2,3]*

[1]*School of Agriculture and Aquaculture, Tra Vinh University, Vinh Long 85000, Vietnam*
[2]*Functional Genomics Research Center, NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City 72820, Vietnam*
[3]*Center for Hi-Tech Development, Nguyen Tat Thanh University, Saigon Hi-Tech Park, Ho Chi Minh City, Vietnam*

ARTICLE INFO

Article history:
Received July 23, 2025
Received in revised form September 18, 2025
Accepted October 17, 2025
Available Online October 24, 2025

KEYWORDS:
comparative genomics,
interspecies,
intraspecies,
plastome evolution

ABSTRACT

The chloroplast genome (cpDNA) plays an important role in the evolutionary history of various plant species. However, most records of cpDNA focused on a representative of a plant taxon. How are interspecies and intraspecies cpDNA evolutionary patterns? To provide hints for this question, we conducted comparative genomic analyses of genome size, structure, gene content, and sequence variations among examined Calophyllaceae interspecific and intraspecific cpDNAs. The results indicated a conservation of genome structure and gene content in both interspecies and intraspecies groups. Within the intraspecies group, *Calophyllum inophyllum* had a stable genome size and high pairwise identity (99.998%) compared to a variable genome size (from 161,444 bp to 161,473 bp) and pairwise identity (from 99.7% to 99.98%) of *Mesua ferrea*. Consequently, *M. ferrea* cpDNA has higher single-nucleotide polymorphism sites and indels compared to *C. inophyllum*. Otherwise, interspecies cpDNAs exhibited a diversity of genome size and different sequence variation rates between *Calophyllum* and *Kielmeyera* species. Comparative genomic analyses revealed conserved and diverse patterns during the evolutionary history of interspecies and intraspecies Calophyllaceae cpDNAs.

1. Introduction

The chloroplast genome is an essential part of angiosperms, which contains genes responsible for performing photosynthesis (Daniell *et al*. 2016; Dobrogojski *et al*. 2020). Additionally, chloroplast genomes are useful genomic resources to trace evolutionary history, to explore population genetics, and to develop molecular markers (Song *et al*. 2023; Yue *et al*. 2023). Previously, studies of the chloroplast genome focused on one individual, which was treated as a representative species in the family and orders (Do *et al*. 2013; Henriquez *et al*. 2014; Downie and Jansen 2015; Givnish *et al*. 2016). Therefore, most of the previous chloroplast genome studies provided information about the interspecies evolution, which revealed a high conservation of chloroplast genomes regarding quadripartite structure (including a large single copy (LSC), a small single copy (SSC), and two inverted repeat (IR) regions) and gene content (including 79 protein-coding genes, 30 tRNA genes, and four rRNA genes). However, some exceptions were observed in the chloroplast genome, such as the lack of one inverted repeat region in Fabaceae (Qin *et al*. 2025) and the absence of the SSC region in *Asarum* species (Lim *et al*. 2018; Wang *et al*. 2021; Choi *et al*.

---

*Corresponding Author
E-mail Address: dhdkhoa@ntt.edu.vn

2024). Additionally, gene loss and duplication events were reported in different plant families (Do *et al*. 2014; Yao *et al*. 2015; Xu *et al*. 2015). These results indicated specific changes in the chloroplast genomes across land plants.

Recently, the advancement of sequencing technology triggered the rise of pan-chloroplast genomes of plants (Magdy *et al*. 2019; Go *et al*. 2024; Wang *et al*. 2024). Specifically, 321 complete chloroplast genomes of five species and two varieties of pepper were compared to identify variable sites, which were successfully used to develop molecular markers (Magdy *et al*. 2019). Another pan-chloroplast genome study of 95 *Hibiscus syriacus* revealed the conservation of genome structure and gene content and proposed the *trnS-psbZ* region as a potential sequence for distinguishing *H. syriacus* cultivars (Go *et al*. 2024). A comparative genomic study of the 322 *Prunus mume* chloroplast genomes located in China and Japan demonstrated a classification of six lineages, which are useful for further genomic studies of *P. mume* (Wang *et al*. 2024). The increasing number of pan-chloroplast genome studies will provide new insights into the evolution of plants.

Most recent pan-chloroplast genome studies focused on ornamental, industrial, and agricultural plants. The results indicated a diversity pattern of chloroplast genome sequences, although gene content and genome structure were conserved. Therefore, a question remained for wild plants: whether interspecies and intraspecies chloroplast genomes had a diversity or conservation pattern? We hypothesize that the chloroplast genomes of interspecies and intraspecies Calophyllaceae exhibit both conserved and diverse patterns during evolutionary history. To deal with this issue, we employed ten complete chloroplast genomes of seven species of Calophyllaceae, representing both interspecies and intraspecies groups, to conduct comparative analysis regarding genome size, genome structure, gene content, and sequence variation. The number of samples in the current study is small; however, the outcomes of this study provide initial information about evolutionary patterns of interspecies and intraspecies chloroplast genomes, which are essential for further studies examining population genetics and molecular marker development.

2. Materials and Methods

2.1. Data Acquisition

Previously, chloroplast genomes of Calophyllaceae were reported, of which the chloroplast genome of *Calophyllum brasiliense* (MW853786) was a partial sequence, so it was not used in the current study (Trad *et al*. 2021). Additionally, in the current study, we focused on interspecies and intraspecies comparisons within the genera of Callophyllaceae. However, only the complete chloroplast genomes of *Caraipa heterocarpa* (MW853787) and *Mahurea exstipulata* (MW853790) were available in the *Caraipa* and *Mahurea* genera, respectively, and thus were not employed for comparative analyses. In a previous study, we reported the complete chloroplast genome of *Calophyllum inophyllum* but not an in-depth comparative analysis (Huynh *et al*. 2024). Therefore, in the current study, we used our reported chloroplast genome and an available sequence of *C. inophyllum* (OL679531) for further analyses. After screening the complete chloroplast genomes that are suitable for interspecies and intraspecies within genera, ten complete chloroplast genomes of seven Calophyllaceae species were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/nuccore) under GenBank format (Table 1). The dataset includes three records of *Mesua ferrea*, two chloroplast genomes of *C. inophyllum*, four species of the *Calophyllum genus*, and two members of the *Kielmeyera* genus (Table 1).

2.2. Genome Feature Characterization, Comparative Genomic Analysis, and Variant Discovery

Geneious Prime is a bioinformatics software that integrates various tools for sequence analysis (https://www.geneious.com/). Therefore, in the current study, we employed various functions related to genome size, genome structure, gene content, and sequence variation for analyzing the data. Specifically, for characterizing the basic features, ten chloroplast genome sequences were imported to Geneious Prime v2024.1 to extract the information of genome size, genome structure (including four regions of large single copy -LSC, small single copy-SSC, and two inverted repeats-IRs), gene content, and GC (Guanine and Cytosine) content. For comparative genomic analysis, the complete chloroplast

Table 1. List of examined Calophyllaceae species and their chloroplast genomes features

Species	Accession number	Total length (bp) / %GC	LSC length (bp) / %GC	IR length (bp) / %GC	SSC length (bp)/ %GC	Unique protein-coding genes/ tRNAs/rRNAs
Mesua ferrea	MK995181	161,473 / 36.4	88,763 / 34	27,614 / 42.1	17,482 / 30.6	79/30/04
	MN052680	161,470 / 36.4	88,760 / 34	27,614 / 42.1	17,482 / 30.6	79/30/04
	MT621164	161,444 / 36.4	88,747 / 34	27,614 / 42.1	17,470 / 30.6	79/30/04
Calophyllum membranaceum	MW801233	160,820 / 36.4	88,913 / 34	27,165 / 42.3	17,577 / 30.6	79/30/04
Calophyllum polyanthum	ON881640	160,750 / 36.5	88,971 / 34	27,195 / 42.3	17,389 / 30.6	79/30/04
Calophyllum soulattri	OP573228	160,381 / 36.4	88,680 / 34	27,624 / 42.0	17,453 / 30.6	79/30/04
Calophyllum inophyllum	OL679531	161,169 / 36.4	89,049 / 33.9	27,345 / 42.2	17,430 / 30.6	79/30/04
Calophyllum inophyllum	PP808711	161,169 / 36.4	89,049 / 33.9	27,345 / 42.2	17,430 / 30.6	79/30/04
Kielmeyera coriacea	MW853789	159,470 / 36.6	88,263 / 34.3	26,490 / 42.7	18,219 / 30.6	79/30/04
Kielmeyera appariciana	MW853788	158,986 / 36.6	87,648 / 34.2	26,519 / 42.7	18,300 / 30.5	79/30/04

genomes were aligned using MAUVE embedded in Geneious Prime v2024.1 with default settings (Darling *et al*. 2004). The aligned sequences were used to calculate the pairwise identity among surveyed intraspecies and interspecies members of Calophyllaceae using Geneious Prime v2024.1. To locate the sites of single-nucleotide polymorphism (SNP) and insertion and deletion (indels) among the examined Calophyllaceae taxa, the function of "Find Variants/SNPS" of Geneious Prime v2024.1 was used with the options of "1" for Minimum Coverage, "0.25" for Minimum Variant Frequency, and other default settings.

3. Results

3.1. Intraspecies and Interspecies Chloroplast Genome Features

Among surveyed intraspecies and interspecies chloroplast genomes of Calophylaceae, genome size varied from 161,444 bp to 161,473 bp in *M. ferrea* individuals, from 160,381 bp to 161,169 bp in *Calophyllum* species, and from 158,986 bp to 159,470 bp in *Kielmeyera* species (Table 1). Consequently, the lengths of LSC, SSC, and IR regions also fluctuated among the examined Calophyllaceae. In contrast to genome size variation, the total GC content of chloroplast genomes was quite stable, except for *Calophyllum polyanthum*, of which the GC content was 0.1 % higher than that of other *Calophyllum* species. Similarly, the GC content of LSC, SSC, and IR regions remained unchanged in the intraspecies chloroplast genome, whereas a slight variation was found in the interspecies chloroplast genome. For example, there was a 0.1 % difference in the GC content of LSC and SSC regions of the *Kielmeyera* species. Among *Calophyllum* species, the GC content of the IR region varied from 42% to 42.3%, while that of the LSC region altered by 0.1% (Table 1).

In contrast to the variation of genome size and GC content, the genome structure of interspecies and intraspecies chloroplast genomes of Calophyllaceae was stable, which had a quadripartite structure including a large single copy, a small single copy, and two inverted repeat regions. Consequently, the number of core genes was the same among examined species, including 79 protein-coding genes, 30 tRNA genes, and four rRNA genes (Table 1). However, there were minor changes at the junctions between LSC, SSC, and IR regions (Figure 1). For example, the LSC/SSC/IR junctions were the same in *C. inophyllum*. A similar trend was observed in *M. ferrea*, except for one record with a longer length (86 bp) from the junction site to the end of the *ndhF* gene. Among *Calophyllum* and *Kielmeyera* species, there was a shift of LSC/IRB junctions from the intergenic spacer between *rps19* and *rpl2* in *C. polyanthum* to the coding region of *rps19* in two *Kielmeyera* species, *C. inophyllum*, and *C. membranaceum*, and to the intergenic spacer between *rps19* and *rpl22* in *C. soulattri*. The SSC/IR junctions are located stably in the *ycf1* coding region. Still, the length from junction site to the start codon of *ycf1* varies among surveyed species (i.e., 1117 bp in *Kielmeyera* species, from 1736 bp to 1909 bp in *Calophyllum taxa*, and 1908 bp in *Mesua ferrea*). Another difference in the SSC/IR boundary is the space from the junction site to the end of the *ndhF* gene, which was adjacent to *ycf1* and *ndhF* in *K. appariciana* and a space from 3 bp to 86 bp in other examined species. A similar pattern was also

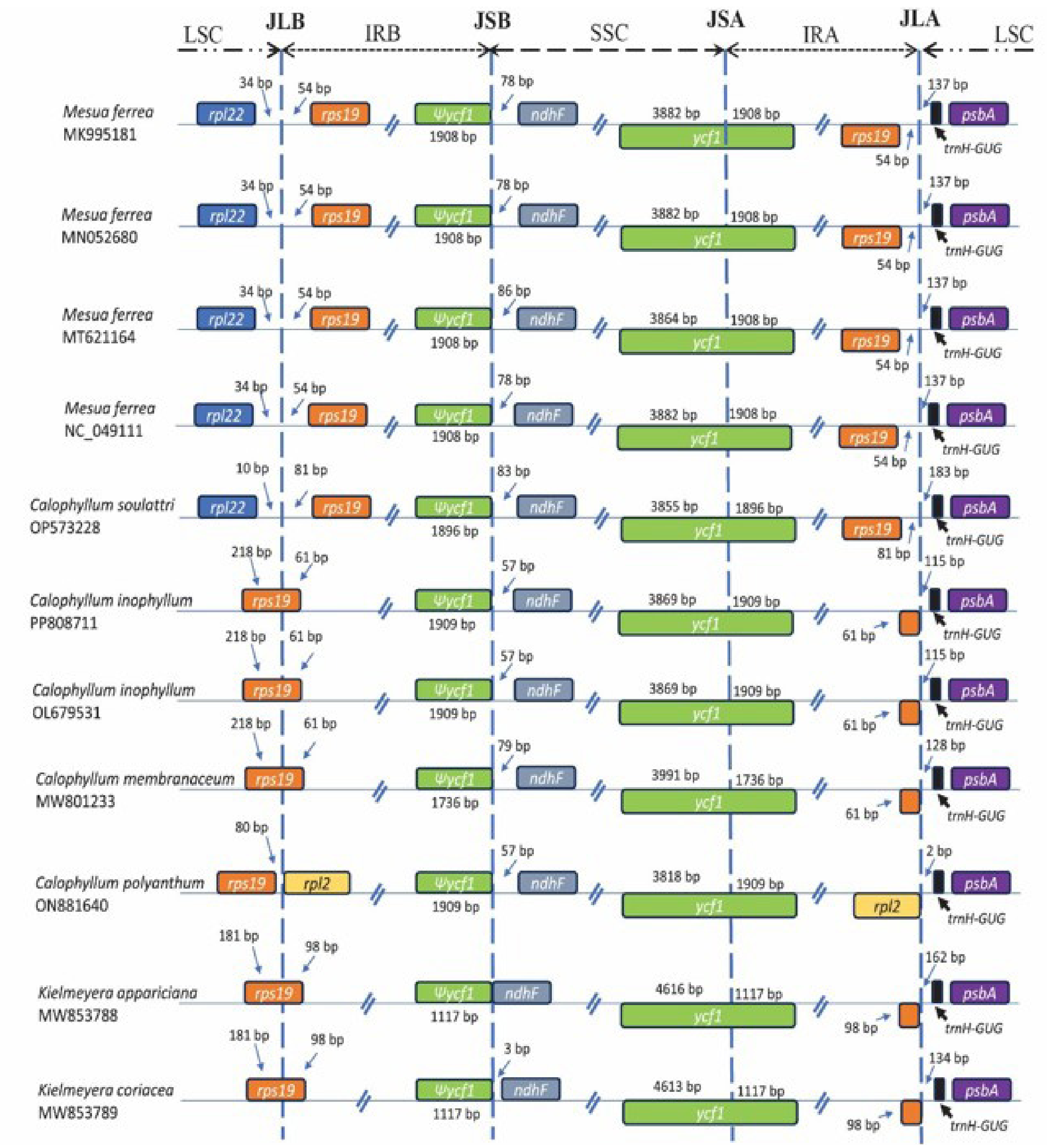

Figure 1. Comparison of junctions between LSC, SSC, and IR regions of Calophyllaceae chloroplast genomes. JLB: junction between LSC and IRB regions. JSB: junction between SSC and IRB regions. JSA: junction between SSC and IRA regions. JLA: junction between LSC and IRA regions. The vertical dashed line indicated the junction sites. The lines represent the sequences and regions and do not show the exact lengths of the chloroplast genomes

found in the LSC/IRA junction, of which the length from the junction site to *trnH-GUG* ranged from 2 bp to 137 bp among surveyed Calophyllaceae species (Figure 1).

3.2. Variations Among Examined Chloroplast Genomes of Calophyllaceae

The similarity of chloroplast genomes among intraspecific Calophyllaceae varied from 99.7% to 99.998% (Table 2). Within *C. inophyllum*, the pairwise identity is 99.998 %. The pairwise identity of interspecies chloroplast genomes was also high but lower than 99.7 % as observed in the intraspecies group (Table 3). Notably, the similarity of *Kielmeyera* and *Calophyllum* species was lower than 99%, which was 98.499% between *K. appariciana* and *K. coriacea*, 98.988% between *C. soulattri* and *C. membranaceum*, and 98.916% between *C. soulattri* and *C. inophyllum*.

Further analysis of SNPs and indels revealed differences between intraspecific and interspecific chloroplast genomes (Table 4). Specifically, the number of SNPs in intraspecies chloroplast genomes (i.e., 81 sites in *M. ferrea* and 3 sites in *C. inophyllum*) was significantly lower than that of interspecies chloroplast genomes of *Callophyllum* (611 sites) and *Kielmeyera* species (287 sites). A higher number of indels was also observed in *Calophyllum* (213 sites) and *Kielmeyera* interspecies group (158 sites) compared to 63 indels in *M. ferrea* (Table 4, Table A.1, Appendix). Another remarkable difference was the length of indels among surveyed Calophyllaceae species. In particular, the intraspecies chloroplast genomes had 1-20 bp indels, whereas the lengths of indels in interspecies chloroplast genomes ranged from 1 bp to 457 bp. Within the intraspecies group, *C. inophyllum* had fewer SNPs and no indels in comparison to 81 SNPs and 63 indels in *M. ferrea*. Among the interspecies group, two *Kielmeyera* species had fewer SNPs and indels but longer lengths of indels in comparison to *Calophyllum* species (Table 4, Table A.1, Appendix).

4. Discussion

"Whether interspecies and intraspecies chloroplast genomes of plants had a diversity or conservation pattern?" is not a difficult question because the answer is predictable. Specifically, some plants will exhibit a conservation trend or a diversity pattern, whereas others might have both patterns of chloroplast genomes. However, which family or which genus possessed these patterns in their chloroplast genomes? The answer to this question remains unknown, and the diversity or conservation pattern might affect further applied research

Table 2. Pairwise identity of intraspecific chloroplast genomes of Calophyllaceae

Species	*Calophyllum inophyllum* (PP808711) (%)	*Mesua ferra* (MK995181) (%)	*Mesua ferra* (MN052680) (%)	*Mesua ferra* (MT621164) (%)	*Mesua ferra* (NC_049111) (%)
Calophyllum inophyllum (OL679531)	99.998	N/A	N/A	N/A	N/A
Mesua ferra (MK995181)	N/A	100	99.704	99.952	99.972
Mesua ferra (MN052680)	N/A	99.704	100	99.700	99.720
Mesua ferra (MT621164)	N/A	99.952	99.700	100	99.980
Mesua ferra (NC_049111)	N/A	99.972	99.720	99.980	100

Table 3. Pairwise identity of interspecific chloroplast genomes of Calophyllaceae

Species	*Kielmyera appariciana* (MW853788) (%)	*Kielmyera coriaceae* (MW853789) (%)	*Calophyllum membranaceum* (MW801233) (%)	*Calophyllum inophyllum* (OL679531) (%)	*Calophyllum polyanthum* (ON881640) (%)	*Calophyllum soulattri* (OP573228) (%)
Kielmyera appariciana (MW853788)	100	98.499	N/A	N/A	N/A	N/A
Kielmyera coriaceae (MW853789)	98.499	100	N/A	N/A	N/A	N/A
Calophyllum membranaceum (MW801233)	N/A	N/A	100	99.117	99.591	98.988
Calophyllum inophyllum (OL679531)	N/A	N/A	99.117	100	99.215	98.916
Calophyllum polyanthum (ON881640)	N/A	N/A	99.591	99.125	100	99.058
Calophyllum soulattri (OP573228)	N/A	N/A	98.988	98.916	99.058	100

Table 4. Summary of SNP and Indels quantities and size of indels in the chloroplast genomes of the examined Calophyllaceae species

Group	Number of SNPs	Number of indels	Length of indels (bp)
Mesua intraspecies	81	63	1-20
Calophyllum intraspecies	3	0	0
Calophyllum interspecies	611	213	1-151
Kielmeyera interspecies	287	158	1-457

based on chloroplast genomes. Therefore, the question seems simple, but the answer plays an important role in further applied genomic studies of plants.

In the current study, we found a conservation of typical quadripartite structure among interspecies and intraspecies Calophyllaceae chloroplast genomes. As a result, the gene content of chloroplast genomes is quite stable, which contains 79 protein-coding genes, 30 tRNA genes, and four rRNA genes. Previous studies revealed a stable structure of the chloroplast genome with a large single copy, a small single copy, and two inverted repeat regions (Xu *et al*. 2015; Smith 2017; Dobrogojski *et al.* 2020). However, the tripartite structure was also found in Fabaceae and Aristolochiaceae, caused by the lack of IR and SSC regions, respectively (Lim *et al*. 2018; Wang *et al*. 2021; Choi *et al*. 2024; Qin *et al*. 2025). Additionally, there were reports about gene loss (i.e., loss of *infA* and *rps16* across angiosperms) and gene duplication (such as triplication of *trnI-CAU* in Melanthiaceae and duplication of *trnfM-CAU* in Actinidiaceae) in chloroplast genomes (Millen *et al*. 2001; Do *et al*. 2014; Yao *et al*. 2015; Schwarz *et al*. 2015; Mohanta *et al*. 2020; Alqahtani and Jansen 2021). Another remarkable event regarding gene content was the pseudogenization and deletion of *ndh* genes in Orchidaceae resulting from the transition from photosynthetic to non-photosynthetic lifestyles (Lin *et al*. 2015, 2017; Kim *et al*. 2023). Among non-photosynthetic plants, the genome structure and gene content of the plastid genome changed dramatically (Schelkunov *et al*. 2015; Wicke and Naumann 2018; Wen *et al*. 2022; Goedderz *et al*. 2024). For example, the plastid genome of *Didymoplexis pallens* had a quadripartite structure, whereas those of Gastrodia species only remained circular (Wen *et al.* 2022). Another report of plastid genomes in the *Epipogium* genus revealed a quadripartite and tripartite structure in *E. roseum* and *E. aphyllum*, respectively (Schelkunov *et al.* 2015). Besides the alteration of the genome structure, the gene content of non-photosynthetic organisms decreased extremely. For instance, only 21 tRNA, 4 rRNA, and 36 protein-coding genes were annotated in the plastid genome of *Danxiaorchis singchiana* (Orchidaceae) (Lee *et al.* 2020). A more severe reduction of gene content was found in *E. roseum* and *E. aphyllum*, of which 29 and 27 unique genes were observed in their plastid genome (Schelkunov *et al*. 2015). Different comparative studies of plastid genomes among non-photosynthetic plants have been conducted; however, they focused on the interspecies level, which exhibited significant changes in genome structure and gene content. A recent genomic study of the holomycotrophic orchid, *Neottia listeroides*, revealed a quadripartite structure but a difference in genome size between samples collected in Tibet, China (94,499 bp) and those from Yunnan, China (from 107,116 bp to 110,855 bp) (Shao *et al.* 2023). Furthermore, the gene content of *N. listeroides* was stable with 35 protein-coding genes, eight rRNA genes, and 38 tRNA genes. These results revealed that non-photosynthetic plants also exhibited both diverse and conserved patterns of plastid genomes, although they had higher selection pressure than their photosynthetic counterparts. Therefore, further research on intraspecies plastid genomes of non-photosynthetic plants is needed to uncover the hidden tracks of evolution.

Although genome structure and gene content were quite stable among interspecific and intraspecific chloroplast genomes, the genome size varied across plants. For example, comparative analysis of 14 species, including mycoheterotrophic and photosynthetic plants of Vanilloideae and Apostasioideae, revealed a significant difference of genome size, which varied from 70,498 bp to 162,174 bp (Kim *et al*. 2023). The remarkable variation in genome size between photosynthetic and non-photosynthetic plants may be attributed to deletions within the plastid genome of non-photosynthetic plants. However, among autotrophic plants, the genome size of interspecific chloroplast genomes was also altered, resulting from the IR expansion. For instance, among eight examined *Polygonatum* species, the boundary between IRb and LSC regions shifted within the *rpl2-rpl22* region, resulting in the variation of IR region length from 25,008 bp to 26,451 bp (Yan *et al.* 2023). Similarly, the shift of the IR/SSC junction was observed in five *Dipodium* species chloroplast genomes, of which the total length ranged from 142,949 bp to 152,956 bp (Goedderz *et al*. 2024). However, variation of LSC/SSC/IR junctions occurred at different rates (Jost and Wanke 2024). In particular, among *Piperales* species, the chloroplast genome of Saururaceae had stable junctions of LSC/SSC/IR regions, whereas those of *Asarum* (Aristolochiaceae) species exhibited extreme changes of the boundaries. In the current study, the variation of IR/LSC/SSC junctions was more severe in *Calophyllum* species than in *Kielmeyera* species (Figure 1). Consequently, the genome sizes of *Calophyllum* chloroplast genomes altered from 160,381 bp to 161,169 bp, whereas those of *Kielmeyera* species were from 158,986 bp to 159,470 bp (Table 1). On the other hand, the intraspecific chloroplast genome of Calophyllaceae possessed steady junctions (Figure 1). Therefore, the genome size of the intraspecific chloroplast genome should be the same (i.e., *C. inophyllum* chloroplast genome) or variable (i.e., *M. ferrea* chloroplast genome). Another genomic study revealed that variation of SSC/IRb junctions was observed in *Paphiopedilum barbigerum* chloroplast genomes (Guo *et al*. 2021). A recent study examining pan-chloroplast genomes of *Prunus mume* distributed in China and Japan revealed both constant and variation in genome size, whereas the genome structure and gene content were constant (Wang *et al*. 2024). Another chloroplast genome-based population genetics study of a living fossil plant, *Ginkgo biloba*, also exhibited the patterns of conservation and diversity of chloroplast genomes regarding genome structure, genome size, and gene content (Nie *et al*. 2025). These results indicated the variety of genome size among interspecific and intraspecific chloroplast genomes. Although the same genome size was found among intraspecific chloroplast genomes, there has not been a record of the same size of interspecific chloroplast genomes within a plant family. Therefore, more chloroplast genome data of interspecific samples should be included in further comparative genomic studies.

For sequence variation analysis, the number of SNPs and indels has been quantified in interspecific and intraspecific chloroplast genomes of different plant families. For example, there were 118 SNPs (composed of 45 transitions and 73 transversions) and 130 indels (including 47 non-repeat-, 22 repeat-, and 61 SSR indels) in 65 plastomes of *Ziziphus jujuba* (Hu *et al*. 2022). Similarly, 95 *H. syriacus* chloroplast genomes contained 193 SNPs and 61 indels, whereas 322 *P. mume* chloroplast genomes had 455 SNPs and 201 indels (Go *et al*. 2024; Wang *et al*. 2024). Among interspecies chloroplast genomes, the quantity of SNPs and indels differed from family to family. For instance, there were 1476 SNPs and 322 indels in the nine *Camellia* species chloroplast genomes (Chen *et al*. 2023). In Araliaceae, seven *Panax* species had 1783 SNPs in the chloroplast genomes (Nguyen *et al*. 2020). Most previous studies examining sequence variation focused on the quantification and characterization of SNPs and indels within an interspecies or intraspecies group. A recent study about the diversification of peanuts revealed variation of SNPs and indels in the intraspecific and interspecific chloroplast genome (Zheng *et al*. 2024). These previous results indicated a variable trend of interspecies and intraspecies chloroplast genomes. In the current study, we also found the different quantities of SNPs and indels among intraspecific and interspecific chloroplast genomes of Calophyllaceae, which might reflect the evolutionary patterns of each group. Specifically, the intraspecies chloroplast genomes exhibited a conservation trend in *C. inophyllum* and a diversity in *M. ferrea*. Meanwhile, only a diverse pattern was found in the interspecific chloroplast genomes of Callophyllaceae. These results suggested co-existence of both patterns in intraspecies chloroplast genomes, whereas this phenomenon in interspecies chloroplast genomes needs further verification.

Obviously, the number of complete Calophyllaceae chloroplast genomes in the current study is small. It might not be sufficient to confirm the entire evolutionary history of interspecies and intraspecies chloroplast genomes. Additionally, Calophyllaceae might not contain chloroplast genomes that exhibit all possibilities of evolutionary trends. Furthermore, genome size, genome structure, gene content, and sequence variation of chloroplast genomes might not reflect all aspects of evolutionary history. However, the current results from Calophyllaceae and previous findings in different plant species shed new light on the evolution of the chloroplast genome at the intraspecific and interspecific levels. Therefore, further genomic studies should enlarge the sample size and focus on comparative analysis between species and genera of the family. Additionally, large plant families such as Asteraceae, Orchidaceae, and Fabaceae are good candidates for exploring the evolutionary trends of interspecies and intraspecies chloroplast genomes.

In conclusion, the previous and current genomic data provided the possibility of conserved and diverse patterns of interspecies and intraspecies chloroplast genomes regarding genome structure, gene content, genome size, and sequence variation among plant species. We demonstrated the co-existence of diversity and conservation of intraspecific chloroplast genomes (Figure 2). However, the conservation of interspecies chloroplast genomes needs more evidence to be confirmed. Furthermore, plastid genomes of non-photosynthetic plants might contain interesting clues for tracing these two opposite patterns of evolution. The selective pressure of plastid genes should be included in further comparative genomic studies to fulfill the overall evolution of organelles in plants. Finally, we demonstrate a need for pan chloroplast genome studies not only for

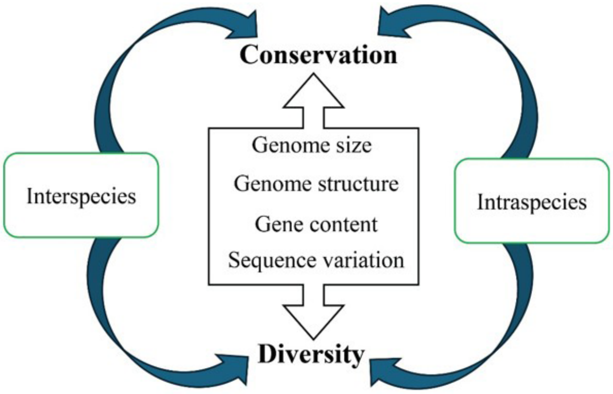

Figure 2. Proposed schematic diagram of evolutionary patterns in intraspecific and interspecific chloroplast genomes of plants

valuable plants but also for wild plants that still hide different aspects of evolution.

Acknowledgements

We acknowledge Tra Vinh University for providing the opportunity to carry out this work. We acknowledge Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam, for supporting this study, especially with high-performing computing devices.

References

Alqahtani, A.A., Jansen, R.K., 2021. The evolutionary fate of rpl32 and rps16 losses in the *Euphorbia schimperi* (Euphorbiaceae) plastome. *Sci. Rep*. 11,7466. https://doi.org/10.1038/s41598-021-86820-z

Chen, J., Zhou, K., Hu, X., 2023. Comparisons of chloroplast genome mutations among 13 Samples of oil-tea camellia from South China. *Genes (Basel)*. 14,1083. https://doi.org/10.3390/genes14051083

Choi, Y-R., Kang, E.S., Kim, T.-H., Choi, Y.-R., Kim, H.-J., 2024. Complete chloroplast genome of *Asarum chungbuensis* (C.S. Yook & J.G. Kim) B.U. Oh 2005 (Aristolochiaceae), a Korean endemic species. *Mitochondrial DNA Part B*. 9,1005-1009. https://doi.org/10.1080/23802359.2024.2387262

Daniell, H., Lin, C.S., Yu, M., Chang, W.-J., 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol*. 17,134. https://doi.org/10.1186/s13059-016-1004-2

Darling, A.C.E., Mau, B., Blattner, F.R., Perna, N.T., 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. 14,1394-1403. https://doi.org/10.1101/gr.2289704

Do, H.D.K., Kim, J.S., Kim, J.-H., 2013. Comparative genomics of four Liliales families inferred from the complete chloroplast genome sequence of *Veratrum patulum* O. Loes. (Melanthiaceae). *Gene*. 530, 229-235. https://doi.org/10.1016/j.gene.2013.07.100

Do, H.D.K., Kim, J.S., Kim, J.-H., 2014. A trnI-CAU triplication event in the complete chloroplast genome of *Paris verticillata* M.Bieb. (Melanthiaceae, Liliales). *Genome Biol. Evol*. 6,1699-1706. https://doi.org/10.1093/gbe/evu138

Dobrogojski, J., Adamiec, M., Luciński, R., 2020. The chloroplast genome: a review. Acta Physiol. *Plant*. 42, 98. https://doi.org/10.1007/s11738-020-03089-x

Downie, S.R., Jansen, R.K., 2015. A Comparative analysis of whole plastid genomes from the Apiales: Expansion and contraction of the inverted repeat, mitochondrial to plastid transfer of DNA, and identification of highly divergent noncoding regions. *Syst. Bot*. 40, 336-351. https://doi.org/10.1600/036364415X686620

Givnish, T.J., Zuluaga, A., Marques, I., Lam, V.K.Y., Gomez, M.S., Iles, W.J.D., Ames, M., Spalink, D., Moeller, J.R., Briggs, B.G., Lyon, S.P., Stevenson, D.W., Zomlefer, W., Graham, S.W., 2016. Phylogenomics and historical biogeography of the monocot order Liliales: out of Australia and through Antarctica. *Cladistics*. 32, 581-605. https://doi.org/10.1111/cla.12153

Go, S., Koo, H., Jung, M., Hong, S., Yi, G., Kim, Y.-M., 2024. Pan-chloroplast genomes for accession-specific marker development in *Hibiscus syriacus*. *Sci. Data*. 11, 246. https://doi.org/10.1038/s41597-024-03077-7

Goedderz, S., Clements, M.A., Bbent, S.J., Nicholls, J.A., Patel, V.S., Crayn, D.M., Schluter P.M., Nargar K., 2024. Plastid phylogenomics reveals evolutionary relationships in the mycoheterotrophic orchid genus Dipodium and provides insights into plastid gene degeneration. *Front. Plant. Sci*. 15, 1388537. https://doi.org/10.3389/fpls.2024.1388537

Guo, Y-Y., Yang, J.-X., Bai, M.-Z., Zhang, G.-Q., Liu, Z.-J., 2021. The chloroplast genome evolution of *Venus slipper* (Paphiopedilum): IR expansion, SSC contraction, and highly rearranged SSC regions. *BMC Plant Biol*. 21, 248. https://doi.org/10.1186/s12870-021-03053-y

Henriquez, C.L., Arias, T., Pires, J.C., Croat, T.B., Schaal, B.A., 2014. Phylogenomics of the plant family Araceae. *Mol. Phylogenet. Evol*. 75, 91-102. https://doi.org/10.1016/j.ympev.2014.02.017

Hu, G., Wu, Y., Guo, C., Lu, D., Dong, N., Chen, B., Qiao, Y., Zhang, Y., Pan, Q., 2022. Haplotype analysis of chloroplast genomes for Jujube breeding. *Front. Plant. Sci*. 13, 841767. https://doi.org/10.3389/fpls.2022.841767

Huynh, D.T.P., Truong, T.T.H., Tran, T.T.H., Do, H.D.K., Nguyen, N.N., 2024. Sequencing and characterizing the chloroplast genome of *Calophyllum inophyllum* L. and its potential for sustainable agricultural development. *Tra Vinh University Journal of Science*. 14, 44-49.

Jost, M., Wanke, S., 2024. A comparative analysis of plastome evolution in autotrophic Piperales. *Am. J. Bot*. 111, e16300. https://doi.org/10.1002/ajb2.16300

Kim, Y.-K., Cheon, S.-H., Hong, J.-R., Kim, K.-J., 2023. Evolutionary patterns of the chloroplast genome in Vanilloid orchids (Vanilloideae, Orchidaceae). *Int. J. Mol. Sci*. 24, 3808. https://doi.org/10.3390/ijms24043808

Lee, S.Y., Meng, K., Wang, H., Zhou, R., Liao, W., Chen, F., Zhang, S., Fan, Q., 2020. Severe plastid genome size reduction in a Mycoheterotrophic orchid, *Danxiaorchis singchiana*, reveals heavy gene loss and gene relocations. *Plants*. 9, 521. https://doi.org/10.3390/plants9040521

Lim, C.E., Lee, S.-C., So, S., Han, S.-M., Choi, J.-E., Lee, B.-Y., 2018. The complete chloroplast genome sequence of *Asarum sieboldii* Miq. (Aristolochiaceae), a medicinal plant in Korea. *Mitochondrial DNA Part B*. 3, 118-119. https://doi.org/10.1080/23802359.2018.1424577

Lin, C.-S., Chen, J.J.W., Huang, Y.-T., Chan, M.-T., Daniell, H., Chang, W.-J., Hsu, C.-T., Liao, D.-C., Wu, F.-H., Lin, S.-Y., Liao, C.-F., Deyholos, M.K., Wong, G.K.-S., Albert, V.A., Chou, M.-L., Chen, C.-Y., Shih, M.-C., 2015. The location and translocation of ndh genes of chloroplast origin in the Orchidaceae family. *Sci. Rep*. 5, 9040. https://doi.org/10.1038/srep09040

Lin, C., Chen, J.J.W., Chiu, C.-C., Hsiao, H.C.W., Yang, C.-J., Jin, X.-H., Leebens-Mack, J., de Pamphilis, C.W., Huang, Y.-T., Yang, L.-H., Chang, W.-J., Kui, L., Wong, G.K.-S., Hu, J.-M., Wang, W., Shih, M.-C., 2017. Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. *The Plant Journal*. 90, 994-1006. https://doi.org/10.1111/tpj.13525

Magdy, M., Ou, L., Yu, H., Chen, R., Zhou, Y., Hassan, H., Feng, B., Taitano, N., van der Knaap, E., Zou, X., Li, F., Ouyang, B., 2019. Pan-plastome approach empowers the assessment of genetic variation in cultivated *Capsicum* species. *Hortic. Res*. 6, 108. https://doi.org/10.1038/s41438-019-0191-x

Millen, R.S., Olmstead, R.G., Adams, K.L., Palmer, J.D., Lao, N.T., Heggie, L., Kavanagh, T.A., Hibberd, J.M., Gray, J.C., Morden, C.W., Calie, P.J., Jermiin, L.S., Wolfe, K.H., 2001. Many parallel losses of *infA* from chloroplast DNA during Angiosperm evolution with Multiple Independent Transfers to the nucleus. *Plant Cell*. 13, 645-658. https://doi.org/10.1105/tpc.13.3.645

Mohanta, T.K., Mishra, A.K., Khan, A., Hashem, A., Abd_Allah, E.F., Al-Harrasi, A., 2020. Gene oss and evolution of the plastome. *Genes (Basel)*. 11, 1133. https://doi.org/10.3390/genes11101133

Nguyen, V.B., Giang, V.N.L., Waminal, N.E., Park, H.-S., Kim, N.-H., Jang, W., Lee, J., Yang T.-J., 2020. Comprehensive comparative analysis of chloroplast genomes from seven *Panax* species and development of an authentication system based on species-unique single nucleotide polymorphism markers. *J. Ginseng. Res*. 44, 135-144. https://doi.org/10.1016/j.jgr.2018.06.003

Nie, L., Liu, F., Wang, M., Jiang, Z., Kong, J., Tembrock, L.R., Kan, S., Wang, P., Wang, J., Wu, Z., Liu, X., 2025. Plastome data provides new insights into population differentiation and evolution of Ginkgo in the Sichuan Basin of China. *BMC Plant Biol*. 25, 48. https://doi.org/10.1186/s12870-024-05977-7

Qin, Q., Dong, Y., Chen, J., Wang, B., Peng, Y., Zhang, X., Wang, X., Zeng, J., Zhong, G., Zhang, S., Du, X., 2025. Comparative analysis of chloroplast genomes reveals molecular evolution and phylogenetic relationships within the Papilionoideae of Fabaceae. *BMC Plant Biol.* 25, 157. https://doi.org/10.1186/s12870-025-06138-0

Schelkunov, M.I., Shtratnikova, V.Y., Nuraliev, M.S., Selosse, M.-A., Penin, A.A., Logacheva, M.D., 2015. Exploring the limits for reduction of plastid genomes: A case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. *Genome Biol. Evol*. 7, 1179-1191. https://doi.org/10.1093/gbe/evv019

Schwarz, E.N., Ruhlman, T.A., Sabir, J.S.M., Hajrah, N.H., Alharbi, N.S., Al-Malki, A.L., Bailey, C.D., Jansen, R.K., 2015. Plastid genome sequences of legumes reveal parallel inversions and multiple losses of rps16 in papilionoids. *J. Syst. Evol.* 53, 458-468. https://doi.org/10.1111/jse.12179

Shao, B.-Y., Wang M.-Z., Chen, S.-S., Ya, J.-D., Jin, X.-H., 2023. Habitat-related plastome evolution in the mycoheterotrophic *Neottia listeroides* complex (Orchidaceae, Neottieae). *BMC Plant Biol*. 23, 282. https://doi.org/10.1186/s12870-023-04302-y

Smith, D.R., 2017. Evolution: In chloroplast genomes, anything goes. *Current Biology*. 27, R1305-R1307. https://doi.org/10.1016/j.cub.2017.10.049

Song, Y., Li, C., Liu, L., Hu, P., Li, G., Zhao, X., Zhou, H., 2023. The population genomic analyses of chloroplast genomes shed new insights on the complicated ploidy and evolutionary history in Fragaria. *Front. Plant. Sci*. 13, 1065218. https://doi.org/10.3389/fpls.2022.1065218

Trad, R.J., Cabral, F.N., Bittrich, V., da Silva, S.R., do Carmo Estanislau do Amaral, M., 2021. Calophyllaceae plastomes, their structure and insights in relationships within the clusioids. *Sci. Rep*. 11, 20712. https://doi.org/10.1038/s41598-021-99178-z

Wang, G., Bai, X., Cao, P., Han, J., Zhang, Z., 2021. The complete chloroplast genome of *Asarum pulchellum* Hemsl. (Aristolochiaceae: *Asarum* L.). *Mitochondrial DNA Part B*. 6, 3257-3258. https://doi.org/10.1080/23802359.2021.1992317

Wang, J., Kan, J., Wang, J., Yan, X., Li, Y., Soe, T., Tembrock, L.R., Xing, G., Li, S., Wu, Z., Jia, M., 2024. The pan-plastome of *Prunus mume*: insights into Prunus diversity, phylogeny, and domestication history. *Front. Plant. Sci*. 15, 1404071. https://doi.org/10.3389/fpls.2024.1404071

Wen, Y., Qin, Y., Shao, B., Li, J., Ma, C., Liu, Y., Yang, R., Jin, X., 2022. The extremely reduced, diverged and reconfigured plastomes of the largest mycoheterotrophic orchid lineage. *BMC Plant Biol*. 22, 448. https://doi.org/10.1186/s12870-022-03836-x

Wicke, S., Naumann, J., 2018. Molecular evolution of plastid genomes in parasitic flowering plants. In: Chaw, S.-M., Jansen, R.K. (Eds.). *Advances in Botanical Research*, *1st edition*. London: Academic Press. pp. 315-347. https://doi.org/10.1016/bs.abr.2017.11.014

Xu, J-H., Liu, Q., Hu, W., Wang, T., Xue, Q., Messing, J., 2015. Dynamics of chloroplast genomes in green plants. *Genomics*. 106, 221-231. https://doi.org/10.1016/j.ygeno.2015.07.004

Yan, M., Dong, S., Gong, Q., Xu, Q., Ge, Y., 2023. Comparative chloroplast genome analysis of four Polygonatum species insights into DNA barcoding, evolution, and phylogeny. *Sci. Rep*. 13, 16495. https://doi.org/10.1038/s41598-023-43638-1

Yao, X., Tang, P., Li, Z., Li, D., Liu, Y., Huang, H., 2015. The First complete chloroplast genome sequences in Actinidiaceae: Genome structure and comparative analysis. *PLoS One*. 10, e0129347. https://doi.org/10.1371/journal.pone.0129347

Yue, M., Chen, H., Xuan, L., Yang, Y., Chong, X., Li, M., Yu, C., Lu, X., Zhang, F., 2023. Novel molecular markers for Taxodium breeding from the chloroplast genomes of four artificial Taxodium hybrids. *Front. Genet*. 14, 1193023. https://doi.org/10.3389/fgene.2023.1193023

Zheng, Z., Sun, X., Qi, F., Fang, Y., Lin, K., Pavan, S., Huang, B., Dong, W., Du, P., Tian, M., Shi, L., Xu, J., Han, S., Liu, H., Qin, L., Zhang, Z., Dai, X., Miao, L., Zhao, R., Wang, J., Liao, Y., Li, A., Ruan, J., Delvento, C., Cigliano, R.A., Maliepaard, C., Bai, Y., Visser, R.G.F., Zhang, X., 2024. Chloroplast and whole-genome sequencing shed light on the evolutionary history and phenotypic diversification of peanuts. *Nat. Genet*. 56, 1975-1984. https://doi.org/10.1038/s41588-024-01876-7