Bangladesh J. Bot. **54**(3): 677-685, 2025 (September) Special DOI: https://doi.org/10.3329/bjb.v54i30.85104

CHARACTERIZATION OF THE *LAGERSTROEMIA SUPRARETICULATA* CHLOROPLAST GENOME AND COMPARATIVE ANALYSIS OF THE CRAPE MYRTLE

BO QIN[1,2], XIN HUANG[2], LIMING TANG[3], YAOHENG HUANG[2], RIHONG JIANG[2], KAIDAO SUN[2], YIQIAN JU[4] AND GANGMIN ZHANG[1*]

Laboratory of Systematic Evolution and Biogeography of Woody Plants, School of Ecology and Nature Conservation, Beijing Forestry University, Beijing 100083, China

Keywords: *Lagerstroemia suprareticulata*, Chloroplast genome, Codon bias, Repeat sequences, Sequence divergence, Phylogenetic relationship

Abstract

Lagerstroemia suprareticulata, a critically endangered species indigenous to the southwestern region of Guangxi, China, holds considerable potential for various applications. Yet, its genetic blueprint remains largely uncharted. We sequenced, assembled, and dissected the chloroplast genome (CPDNA) of *L. suprareticulata* to delineate its genomic architecture, evolutionary divergence, and its phylogenetic stance within the *Lagerstroemia* genus. This CPDNA spans 152,196 bp with a circular quadripartite structure: 84,027 bp large single-copy (LSC), 16,919 bp single-copy (SSC) region, and two 25,625 bp inverted repeat (IR) regions. Notably, genomic structure, gene content, repeats, IR dynamics, and sequence divergence among five related lythraceous species are largely conserved with slight variations. Six intergenic regions and one protein-coding gene exhibit high diversity, useful as molecular markers for phylogeny and species differentiation.. Phylogenetic assessments employing full CPDNA and coding sequences elucidate that *L. suprareticulata* aligns as the progenitor of *L. glabra* and *L. anhuiensis*. This analysis offers a thorough perspective on the CPDNA of *L. suprareticulata*, thus augmenting the genomic resources available for probing its evolutionary lineage and genetic diversity.

Introduction

The petals of *Lagerstroemia* resemble crumpled yarn, which has led to the common English name “Crape myrtle” (or “Crepe myrtle”). *Lagerstroemia* species are distributed globally, predominantly in tropical and subtropical regions, although some species also adapt to temperate climates. This genus is highly valued as an ornamental tree in various landscapes, and its analysis and research hold significant application value (Gu *et al*. 2016). *L. suprareticulata*, a deciduous tree or shrub in the family Lythraceae, is endemic to the southwestern region of Guangxi, China. It is listed as endangered in the Threatened Species List of China’s Higher Plants (Qin and Zhao 2017). This species flourishes in limestone regions, demonstrating potential for use in the greening of rocky mountains and ecological rehabilitation. Additionally, it bears fragrant blossoms, an uncommon feature within the *Lagerstroemia* genus. Considering its significant application prospects and the escalating risks to its existence, it is essential to prioritize conservation initiatives and genetic research to protect its germplasm resources.

Genomes of chloroplasts have shown significant sequence and structural variation among plant species, which has helped to shed insight on evolutionary relationships across different branches of the tree of life. Many woody flowering plants' CPDNA sequences have been obtained, the annotation and sequencing of *Lagerstroemia* genomes stand out because they help to comprehend their evolutionary histories (Dong *et al*. 2021, He *et al*. 2024).

*Author for correspondence: <gary1967@bjfu.edu.cn>. [1]Laboratory of Systematic Evolution and Biogeography of Woody Plants, School of Ecology and Nature Conservation, Beijing Forestry University, Beijing 100083, China; [2]Guangxi Key Laboratory of Special Non-wood Forest Cultivation and Utilization, Guangxi Laboratory of Forestry, Guangxi Forestry Research Institute, Nanning 530002, China. [3]Guangxi Forestry Industry Group Co. LTD, Nanning 530025, China. [4]College of Landscape Architecture and Forestry, Qingdao Agricultural University, Qingdao 266109, China.

This study combines molecular pylogenetic methods and comparative chloroplast genomics to analyze the CPDNA and its evolutionary relationships in *Lagerstroemia* in great detail. The objectives of the study were to learn more about the CPDNA structure and codon usage in *L. suprareticulata*, to compare and contrast the CPDNA structures of related species and, to figure out where *L. suprareticulata* fits in the evolutionary tree of *Lagerstroemia* and how it relates to other members of the genus.

Materials and Methods

Young leaves of *L. suprareticulata* were collected from karst habitats in Daxin, Chongzuo, Guangxi. Vouchers are stored at the Herbarium of Guangxi Forestry Research Institute (No.: 2023050601); DNA samples are preserved at the Guangxi Key Laboratory of Special Non-Wood Forest Cultivation & Utilization (Nanning). For library preparation, 0.2 μg genomic DNA (modified CTAB extraction) was used. Following NovaSeq 6000 protocols, indexed samples underwent sonication to 350 bp, end-repair, A-tailing, and adapter ligation. PCR products were purified with AMPure XP (Beckman Coulter). Libraries (Agilent 4200 quality check; 1.5 nM via QPCR) were pooled and PE150-sequenced on Illumina platforms.

Raw data were processed with Trimmomatic v0.39 (Bolger *et al.* 2014) to remove adapters/low-quality reads. Corrected reads were assembled via SPAdes v3.14.1 (Prjibelski *et al.* 2020, k-mer = 95). BLASTn (Chen *et al.* 2015) and Exonerate identified matches (1e-10, 70% thresholds); low-coverage fragments were filtered. CPDNA was annotated with PGA v3 (Qu *et al.* 2019); circular maps via OGDRAW.

CodonW v1.4.2 analyzed bias using RSCU: >1 (preference), <1 (rarer), =1 (neutral). REPUter assessed repeats (>90% identity, 30 bp min, Hamming distance=3). MISA (Beier *et al.* 2017) detected SSRs: mono (>10 nt), di (>5 nt), tri (>4 nt), tetra/penta/other (>3 nt); 100 nt spacing. IRscope analyzed IR expansion/contraction across four borders (IRa/LSC, IRb/SSC, IRb/IRb, LSC/IRb) in five species.

L. speciosa, *L. suprareticulata*, *L. fauriei*, *Lawsonia inermi*s, and *Punica granatum* CPDNAs were analyzed via mVISTA (Frazer *et al.* 2004). Nucleotide diversity (Pi) was calculated with DNAsp v6.12.03 (Librado and Rozas 2009, 600 bp window, 200 bp step).

Phylogenetic analysis used 21 CPDNAs (18 *Lagerstroemia*; outgroups: *Rotala rotundifolia*, *Lawsonia inermis*, *Punica granatum*). Geneious v2021.1.1 (Kearse *et al.* 2012) annotated sequences; ClustalW2 (Larkin *et al.* 2007) aligned, trimAI v1.4.1 (Capella-Gutiérrez *et al.* 2009) trimmed. Model Finder (Kalyaanamoorthy *et al.* 2017) selected models; MEGA7.0 (Kumar *et al.* 2016) built ML trees (1,000 bootstraps, TBR swapping).

Results and Discussion

The de novo-assembled cpDNA of *L. suprareticulat*a is 152,196 bp with 36.95% GC. It has a circular quadripartite structure: two IR regions (IRa and IRb), one LSC, and one SSC. IRa and IRb each are 25,625 bp; LSC is 84,027 bp, SSC 16,919 bp. Fig. 1 shows its circular map, with GC contents: 35.08 (LSC), 29.47 (SSC), and 44.44% (IR).

The cpDNA of *L. suprareticulata* encodes 133 genes: 103 protein-coding, 8 ribosomal RNA (rRNA), and 37 transfer RNA (tRNA). Ten protein-coding genes, 4 rRNAs, and 7 tRNAs are duplicated in IR regions. IRa contains the *ycf1* pseudogene, arising from IR expansion. Sixty-three genes are related to self-replication and 43 to photosynthesis. There are three distinct sets of genes: six that deal with ATP synthase (the *atp* group), six that are involved with the cytochrome b/f complex (the *pet* group), and eleven that are involved with NADH dehydrogenase (the *ndh* group). Photosystem I and II rely on genes from the *psa* and *psb* families, respectively. Furthermore,

ribosomal protein synthesis is regulated by 21 genes, some of which belong to the *rps* and *rpl* families. Fourteen genes contain introns, including *petB, ndhA,* and *atpF*. The *rps12* gene is trans-spliced with its 3' end in the IR area and its 5' end in the LSC region, whereas the *clpP* and *ycf3* genes both have two introns (Suppl. Tab. 1).

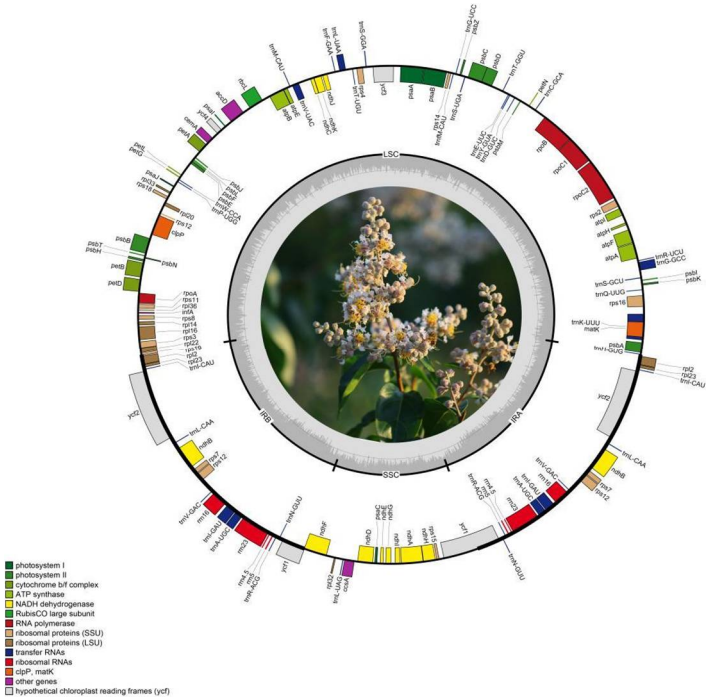

Fig. 1. *L. suprareticulata*'s whole CPDNA gene map. In this diagram, the inner and outer rings of the circle represent genes that are transcribing in opposite directions, respectively. The light gray section inside the inner circle represents AT content, and the dark gray portion represents GC content.

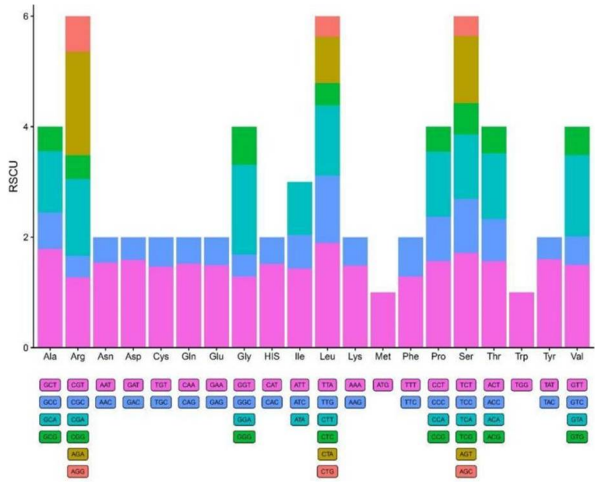

Fig. 2. RSCU of 20 AAs and the stop codon in the complete CPDNA of *L. suprareticulata,* highlighting codon preferences. The histogram colors correspond to the codon types displayed.

As Fig. 2 shows, Leucine (Leu) was most abundant (2873 codons, 10.63%); Cysteine (Cys) least (308 codons, 1.14%). Of 61 codons, TGG (Trp) and ATG (start codon) had RSCU=1. A total of twenty-nine codons exhibited RSCU values greater than 1, indicating a notable codon usage bias, particularly toward A/T-rich codons. The highest RSCU values were observed for TTA (Leu, 1.896), AGA (Arg, 1.871), and GCT (Ala, 1.786).

MISA analysis of *Lagerstroemia*, *Lawsonia*, and *Punica* CPDNAs is shown in the colinear circos plot (Fig. 3). Five CPDNAs contained 194 SSRs, with individual counts ranging from 26 (*L. suprareticulata*) to 61 (*Lawsonia inermis*) (Suppl. Tab. 2). Mononucleotide SSRs dominated (81.48-98.04%), mostly A/T units. Some species lacked dinucleotide and/or trinucleotide SSRs. Dinucleotides were in three species (*L. speciosa*, *Lawsonia inermis*, *Punica granatum*) at 1.96-4.92%, with only AT/AT repeats. Trinucleotides were in three species (*L. suprareticulata*, *L. fauriei*, *L. speciosa*) at 3.45-3.85%, with only AAT/ATT repeats. *L. suprareticulata* had 25 mononucleotides and 1 trinucleotide, no dinucleotide, tetranucleotide, or hexanucleotide SSRs.

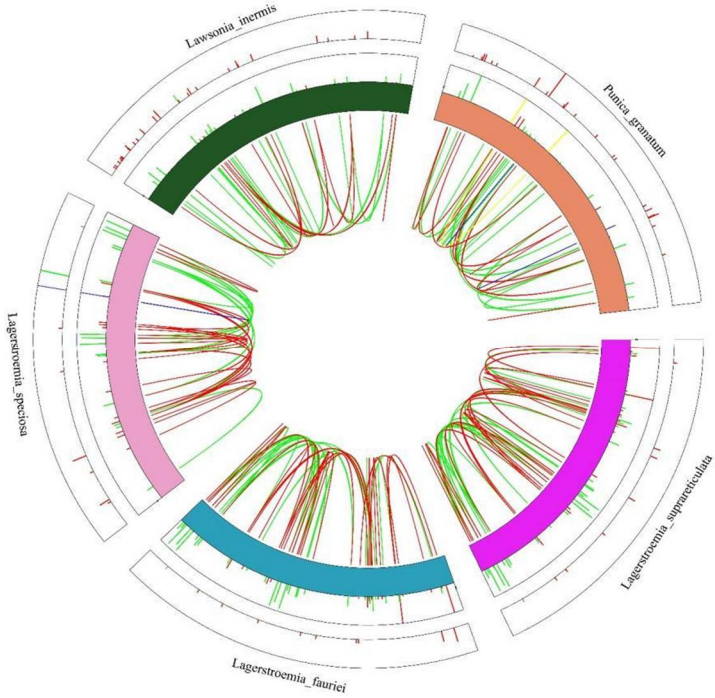

Fig. 3. Analysis of SSRs, dispersed repeats, and collinearity in the CPDNAs of *Lagerstroemia, Lawsonia*, and *Punica* species. The figure consists of concentric circles. The outer circle displays the distribution and repeat numbers (Y-axis) of mononucleotide (red), dinucleotide (yellow), and trinucleotide (blue) SSRs in the CPDNAs. The inner circle histogram illustrates the distribution and lengths of palindromic (green), complement (blue), reverse (yellow), and forward repeat (red) sequences on the Y-axis. The inner circle depicts the repetitive structure of DNA in the CPDNAs of five plant species, highlighting the repetitive relationships of palindromic, complement, reverse, and forward repeats.

Analysis detected 421 long repeats (forward, palindromic, reverse) across five CPDNAs, with lengths 30-25, 973 bp (Fig. 3). Counts: *L. speciosa* (106), *L. suprareticulata* (92), *L. fauriei* (87), *Lawsonia inermis* (76), *Punica granatum* (60). Forward repeats dominated (208), followed by palindromic (196) and reverse (17). *L. suprareticulata* had 47 forward, 44 palindromic, 1 reverse. Most repeats were 30-40 bp, then 41-50 bp.

IRscope analysis revealed notable border zone changes, with varied IR/SC junction types (Fig. 4). *L. suprareticulata* and *L. fauriei* shared one type: *ndhf* at IRb/SSC boundary (JSB). *L. speciosa*, *Lawsonia inermis*, and *Punica granatum* shared another: *ndhf* in SSC. JLB (IRb/LSC boundary) was similar in *L. suprareticulata* and *L. fauriei*, with no significant differences. The latter group showed marked differences: SSC length varied, *ycf1* expanded in three species. *trnH*-IRa distances: *L. speciosa* (0 bp), *Lawsonia inermis* (10 bp), *Punica granatum* (23 bp). *rps* expanded from LSC into IRb by 77 bp (*L. speciosa*), 109 bp (*Lawsonia inermis*), 24 bp (*Punica granatum*). Other species had varying IR contractions (reference: *Punica granatum*).

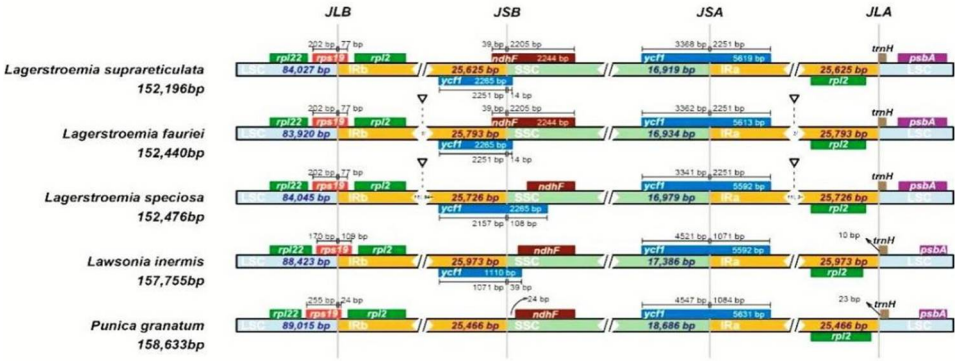

Fig. 4. Genomic comparisons of five chloroplasts concerning their LSC, SSC, and IR regions. Colored boxes show gene locations, and spatial relationships are not shown to scale.

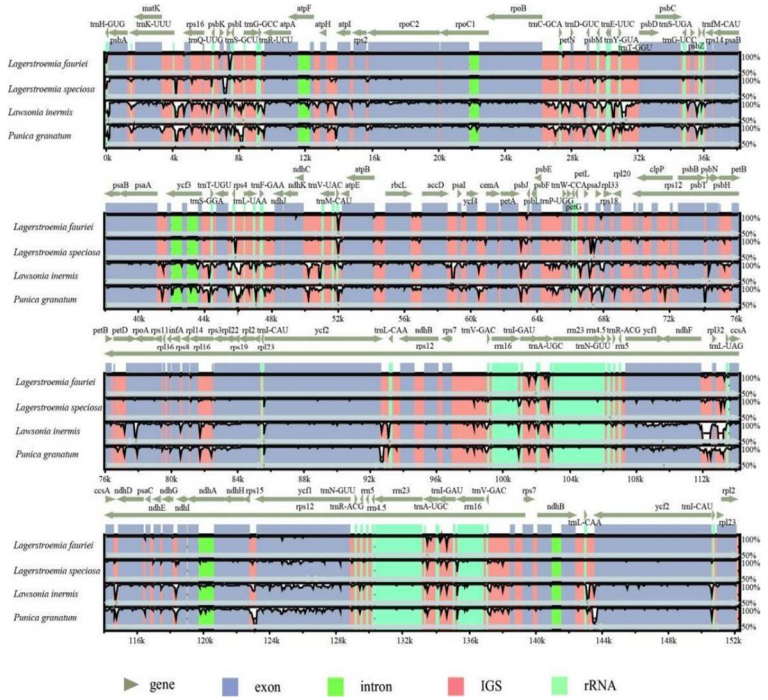

Fig. 5. mVISTA similarity map of five CPDNAs, with *L. suprareticulata* as the reference. The vertical axis represents sequence identity, ranging from 50 to 100%, and genomic regions are color-coded to distinguish between protein-coding, introns, intergenic spaces, and rRNAs.

Fig. 5 shows that the conservation and variability of these genomes were illuminated by mVISTA sequence alignments. Variability was more pronounced in non-coding and single-copy areas, even if gene order and sequence identity were highly conserved. Intergenic regions such as *ndhF-rpl32* and *trnH-GUG-psbA,* showed significant variation, as did coding regions of genes like *rrn16*, *rps16,* and *ycf1*.

Analysis of protein-coding genes and intergenic regions in CPDNAs of five *Lagerstroemia* species revealed intergenic regions were less conserved than protein-coding genes via nucleotide diversity (Pi) analysis. Figure 6a shows high variability in regions with Pi ≥0.004 (max divergence 0.0073). Using *L. suprareticulata* CPDNA as reference, genomic similarity analysis with the other four species also showed greater variation in non-coding vs. coding regions (Fig. 5). Further analysis indicated high-Pi genes were mainly in tRNA, ribosome, and photosystem-I regions, with IR regions relatively conserved (Fig. 6b). Additional highly variable regions (Pi ≥0.07, max divergence 0.016) included *trnH-GUG-psbA, rps16, trnR-UCU-atpA, psaA, trnV-UAC-trnM-CAU, rpl20-rps12,* and *rrn16* across the five CPDNAs.

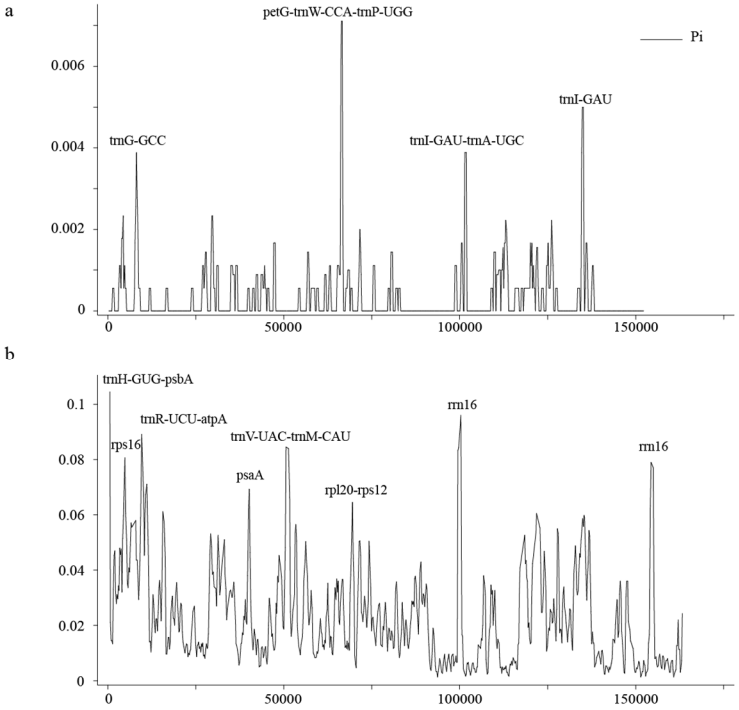

Fig. 6. Examining the Nucleotide Diversity in the CPDNA of *Lagerstroemia* and Other Species: A Comparative Analysis (a) Using a sliding window technique with a 600 bp length and a 200 bp step size, this panel displays the nucleotide diversity across six *Lagerstroemia* species. (b) This figure illustrates the variation in nucleotides found in five distinct species: Punica granatum, Lawsonia inermis, *L. speciosa*, *L. suprareticulata*, and *L. fauriei*. The study uses a 200 bp step size and a 600 bp window length, just like panel (a).

ML phylogenetic trees were constructed from complete CPDNA and all protein-coding genes (CDS) for 22 species (19 *Lagerstroemi*a, with *Rotala rotundifolia*, *Lawsonia inermis, Punica granatum* as outgroups). Both cpDNA and CDS trees strongly supported *Lagerstroemia* monophyly (Fig. 7), with 84.3 and 57.9% of nodes having 100 bootstrap values, respectively, confirming clear interspecific phylogenetic relationships. Both trees confirmed *Lagerstroemia*

divided into three clades (I, II, III), with Clade II and III as sister branches. *L. suprareticulata* was strongly supported as sister to the common ancestor of *L. glabra* and *L. anhuiensis* in both trees. Comparisons revealed conflicting relationships with low bootstrap values at some nodes: *L. fauriei* in Clade II (CPDNA) vs. Clade I (CDS); *L. calyculata-L. loudonii* as sisters (CPDNA) vs. *L. loudonii* sister to *L. calyculata-L. fauriei* ancestor (CDS); in Clade III, *L. guilinensis* diverged first (CPDNA) vs. *L. excelsa first* (CDS), followed by *L. indica* and the other.

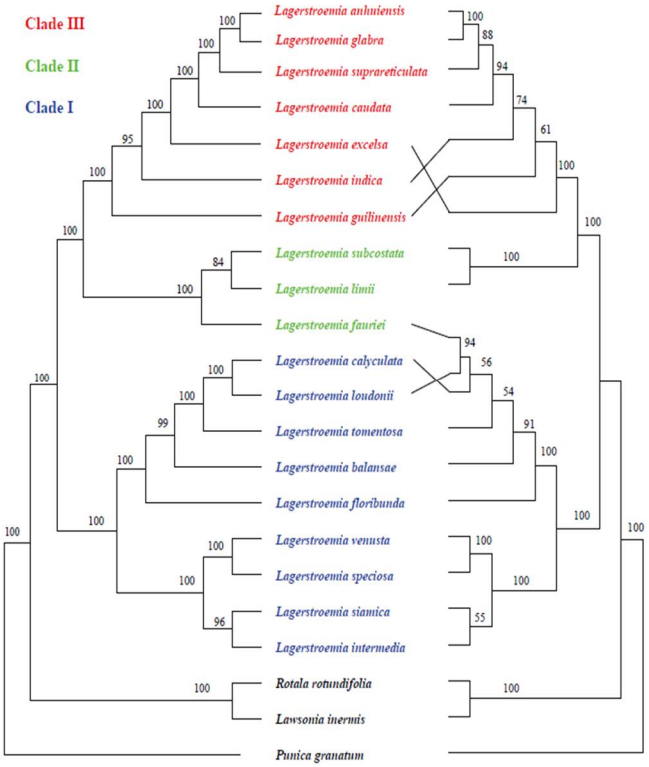

Fig. 7. Phylogenetic Reconstruction of *Lagerstroemia* and Related Species. This figure presents phylogenetic trees constructed via the ML method, based on analyses of the complete CPDNA (left panel) and all 86 protein-coding genes (right panel). The numerals along the branches indicate the ML bootstrap support values for each clade.

Plant CPDNAs typically contain 110-130 genes with conserved sequences and arrangements. The CPDNA of *L. suprareticulata* is somewhat exceptional, with 133 genes-a relatively high count compared to typical plant CPDNAs (He *et al*. 2024). Like most angiosperms, it has a circular quadripartite structure, comprising an LSC region, an SSC region, and two IR regions. Its SSC region (16.9 kb) is relatively small, compared to the 18-20 kb typical of higher plant chloroplasts (Cui *et al*. 2024). Consistent with most plant chloroplasts, its total GC content is similar to that of closely related species, with the IR region having higher GC content (Gu *et al*. 2019).

Repeat sequences underpin CPDNA structural integrity and rearrangement. Their great variation in length and type aids understanding of species phylogeny, biogeography, and population genetics. *L. suprareticulata* has 47 forward and 44 palindromic repeats, with forward repeats predominant across species; the complexity of these repeats is vital for studying CPDNA evolutionary dynamics (Lin *et al*. 2023). Compared to related species, *L. suprareticulata* has fewer, simpler SSRs, predominantly mononucleotide and trinucleotide repeats, which are poised to

facilitate molecular marker development. Prevalent mononucleotides are mainly A/T repeats, less complex than other plant chloroplast SSRs. Notably, a higher A/T proportion-a common *Lagerstroemia* trait significantly influences RSCU distribution (Dong *et al*. 2021).

Dynamic contraction and expansion of IR region boundaries are pivotal evolutionary processes in CPDNAs, markedly affecting genome size. Variability at CPDNA boundaries (JLB, JSB, JSA, JLA) plays a fundamental role in the evolutionary development of specific taxa. Phylogenetic mapping shows *L. fauriei* is more closely related to *L. suprareticulata* than *L. speciosa* is. Notably, *L. fauriei* and *L. suprareticulata* have uniform CPDNA boundaries, while *L. speciosa* differs from *L. suprareticulata*-particularly at JSB, where the *ndhf* gene is further from the boundary. In contrast, outgroup genomes show significant boundary divergences, indicating evolutionary conservation within *Lagerstroemia* CPDNAs.

L. suprareticulata demonstrates high sequence similarity with other Lythraceae members, but some CPDNA regions display significant divergence. mVISTA analysis reveals less pronounced sequence divergence in IR regions, likely due to sequence correction mechanisms during gene replication and transcription (Cui *et al*. 2024). *Lagerstroemia* species differ from other Lythraceae in nucleotide diversity patterns (Fig. 6). Genes with high Pi values, beyond photosynthetic roles, are notably enriched in transfer RNA and ribosomal components. Significant Pi value disparities within *Lagerstroemia* are primarily in the petG gene. This highlights the potential of nucleic acid polymorphism in *Lagerstroemia* for developing molecular markers, aiding molecular identification and marker creation for the genus.

The significance of the CPDNA sequence in the reconstruction of plant phylogenetics is extensively recorded, owing to its elevated mutation rates, maternal inheritance patterns, and lack of recombination. By employing the entire CPDNA sequence, scientists have successfully tackled various phylogenetic questions across both extensive and detailed evolutionary dimensions, thereby deepening the comprehension of intricate evolutionary connections within angiosperms. Phylogenetic analyses employing the complete CPDNA and its protein-coding genes indicate that *Lagerstroemia* can be classified into three separate clades. Notably, the clade that includes *L. suprareticulata* and the common ancestors of *L. glabra* and *L. anhuiensis* is recognized as a sister group, aligning with the conclusions of He *et al*. (2024). Moreover, the discrepancies observed in the topologies between the CPDNA and the CDS trees highlight the variable mutation rates across different genome regions, including coding, non-coding, and spacer areas. Considering the CPDNA's depiction of single-parent inheritance, it is advisable for future studies to integrate nuclear genes to construct a more comprehensive phylogenetic framework. This approach would provide a more balanced view of genetic relationships and evolutionary patterns within the genus.

This research has successfully revealed the sequences and examined the entire CPDNA of *L. suprareticulata*, drawing comparisons with genomes from other species within the Lythraceae family. The essential genetic resources and evolutionary processes, focusing on genome features, codon utilization, repetitive elements, and sequence variation were also investigated. Comparative genomic analyses have demonstrated that although the CPDNA of Lythraceae tends to be conserved, specific regions show significant sequence divergence, which may function as molecular markers for phylogenetic studies. The phylogenetic analysis has provided additional insights into the evolutionary placement of *L. suprareticulata* within its family.

Acknowledgements

The authors are grateful for financial support from the Fundamental Research Funds for Guangxi Forestry Research Institute under Grant No. 202303 and the Guangxi Key Research and Development Program under Grant No. AB25069136.

References

Beier S, Thiel T, Munch T, Scholz U and Mascher M 2017. MISA-web: a web server for microsatellite prediction. Bioinfor. **33**(16): 2583-2585.

Bolger AM, Lohse M and Usadel B 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinfor. **30**(15): 2114-2120.

Capella-Gutiérrez S, Silla-Martínez JM and Gabaldón T 2009. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinfor. **25**(15): 1972-1973.

Chen Y, Ye W, Zhang Y and Xu Y 2015. High speed BLASTN: an accelerated MegaBLAST search tool. Nucl. Acids Res. **43**(16): 7762-7768.

Cui X, Liu K, Li E, Zhang Z and Dong W 2024. Chloroplast genomes evolution and phylogenetic relationships of *Caragana* species. Int. J. Mol. Sci. **25**(12): 6786.

Dong W, Xu C, Liu Y, Shi J, Suo Z. 2021. Chloroplast phylogenomics and divergence times of *Lagerstroemia* (Lythraceae). BMC Genom. **22**(1): 434.

Frazer KA, Pachter L, Poliakov A, Rubin EM and D Innaet 2004. VISTA: computational tools for comparative genomics. Nucl. Acids Res. **32**(Web Server issue): W273-279.

Gu C, Ma L, Wu Z, Chen K and Wang Y 2019. Comparative analyses of chloroplast genomes from 22 Lythraceae species: inferences for phylogenetic relationships and genome evolution within Myrtales. BMC Plant Biol. **19**(1): 281.

Gu C, Tembrock LR, Johnson NG, Simmons MP, Wu Z and Fredericket WR 2016. The Complete plastid genome of *Lagerstroemia fauriei* and loss of rpl2 intron from *Lagerstroemia* (Lythraceae). PLoS One **11**(3): e0150752.

He L, Xu S, Cheng X, Huang H, Dai H, Wang X, Ding Z, Xu M, Gu H and Yan N 2024. Chloroplast genomes in seven *Lagerstroemia* species provide new insights into molecular evolution of photosynthesis genes. Front. Genet. **15**: 1378403.

Kalyaanamoorthy S, Minh BQ, Wong TK and Jermiin LS 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature **14**(6): 587-589.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S and Duran C 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinfor. **28**(12): 1647-1649.

Kumar S, Stecher G and Tamura K 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. **33**(7): 1870-1874.

Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, Mcwilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson J, Gibson TJ and Higgins DG 2007. Clustal W and Clustal X version 2.0. Bioinfor. **23**(21): 2947-2948.

Librado P and Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinfor. **25**(11): 1451-1452.

Lin N, Liu Q, Landis JB, Rana HK, Li Z, Wang H, Sun H and Deng T 2023**.** Staying in situ or shifting range under ongoing climate change: A case of an endemic herb in the Himalaya-Hengduan Mountains across elevational gradients. Div. Dist. **29**(4): 524-542.

Prjibelski AD, Antipov D, Meleshko D, Lapidus A and Korobeynikov A 2020. Using SPAdes de novo assembler. Curr. Prot. Bioinfor. **70**(1): e102.

Qin H and Zhao L 2017. Evaluating the threat status of higher plants in China. Biodiver. Sci. **25**(7): 689.

Qu X, Moore M, Li DZ and Yi T 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Meth. **15**: 50.

Xu Y, Wen J, Su X and Ren Z 2022. Variation among the complete chloroplast genomes of the sumac species rhus chinensis: reannotation and comparative analysis. Genes (Basel) **13**(11).

(Manuscript received on 08 February 2025; revised on 12 October 2025)

Supplementary Table 1. CPDNA gene content.

Function	Group of genes Gene names
ATP synthase	*atpA*, *atpF*, *atpH*, *atpI*, *atpE*, *atpB*
Cytochrome b/f complex	*petN*, *petA*, *petL*, *petG*, *petB*, *petD*
NADH dehydrogenase-like complex	*ndhJ*, *ndhK*, *ndhC*, *ndhB*, *ndhF*, *ndhD*, *ndhE*, *ndhG*, *ndhI*, *ndhA*, *ndhH*
Photosystem I	*psaB*, *psaA*, *psaI*, *psaJ*, *psaC*
Photosystem II	*psbA*, *psbK*, *psbI*, *psbM*, *psbD*, *psbC*, *psbZ*, *psbJ*, *psbL*, *psbF*, *psbE*, *psbB*, *psbT*, *psbN*, *psbH*
Proteins of unknown function	*ycf3*, *ycf4*, *ycf1*, *ycf2*
Ribosomal proteins (SSU)	*rps12*, *rps16*, *rps2*, *rps14*, *rps4*, *rps18*, *rps11*, *rps8*, *rps3*, *rps19*, *rps7*, *rps15*
Ribosomal proteins (LSU)	*rpl33*, *rpl20*, *rpl36*, *rpl14*, *rpl16*, *rpl22*, *rpl2*, *rpl23*, *rpl32*
Ribosomal RNAs	*rrn16*, *rrn23*, *rrn4.5*, *rrn5*
RNA polymerase	*rpoC2*, *rpoC1*, *rpoB*, *rpoA*
Other genes	*matK*, *rbcL*, *accD*, *cemA*, *clpP*, *infA*, *ccsA*
Transfer RNAs	*trnH-GUG*, *trnK-UUU*, *trnQ-UUG*, *trnS-GCU*, *trnG-GCC*, *trnR-UCU*, *trnC-GCA*, *trnD-GUC*, *trnY-GUA*, *trnE-UUC*, *trnT-GGU*, *trnS-UGA*, *trnG-UCC*, *trnfM-CAU*, *trnS-GGA*, *trnT-UGU*, *trnL-UAA*, *trnF-GAA*, *trnV-UAC*, *trnM-CAU*, *trnW-CCA*, *trnP-UGG*, *trnI-CAU*, *trnL-CAA*, *trnV-GAC*, *trnI-GAU*, *trnA-UGC*, *trnR-ACG*, *trnN-GUU*, *trnL-UAG*

Supplementary Table 2. The results of the SSRs analysis for five species.

	A/T	C/G	AT/AT	AAT/ATT
L. Suprareticulata	24 (92.31%)	1 (3.85%)		1 (3.85%)
L. fauriei	27 (93.1%)	1 (3.45%)		1 (3.45%)
L. speciosa	22 (81.48%)	2 (7.41%)	2 (7.41%)	1 (3.7%)
Lawsonia inermis	58 (95.08%)		3 (4.92%)	
Punica granatum	50 (98.04%)		1 (1.96%)	