MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 12, 1642–1647
https://doi.org/10.1080/23802359.2024.2435905

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

The complete chloroplast genome of *Syzygium zeylanicum* (Myrtaceae, Myrtales) and its phylogenetic analysis

Thu-Thao Thi Huynh[a], Minh Trong Quang[b] and Hoang Danh Nguyen[c]

[a]Department of Hematology, Faculty of Medical Laboratory, Hong Bang International University, Ho Chi Minh City, Vietnam; [b]Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam; [c]Functional Genomics Research Center, NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam

ABSTRACT

The complete chloroplast genome of *Syzygium zeylanicum* (L.) DC. 1828 has been sequenced and analyzed for the first time. The *S. zeylanicum* chloroplast genome is 159,445 bp in length, comprised of a large single-copy region (88,034 bp), a small single-copy region (18,455 bp), and a pair of inverted repeat regions (26,478 bp each). The genome encoded 85 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Phylogenetic analysis indicated that *S. zeylanicum* is closely related to *S. acuminatissimum*. This research provides essential genomic data for *S. zeylanicum*, offering valuable resources for future comparative genomics, phylogenetics, and conservation biology studies.

ARTICLE HISTORY
Received 8 August 2024
Accepted 25 November 2024

KEYWORDS
Plastome; genome assembly and annotation; phylogeny; next-generation sequencing

Introduction

The genus *Syzygium* Gaertn. 1788, belonging to the family Myrtaceae, comprises approximately 1200 species of trees and shrubs distributed throughout tropical and subtropical regions (Nigam and Nigam 2012; Thornhill et al. 2015). Many *Syzygium* species are valued for their timber, essential oils, and medicinal properties (Amir Rawa et al. 2022; Uddin et al. 2022). *Syzygium zeylanicum*, known as 'Trâm vỏ đỏ' in Vietnam, is an evergreen tree native to South and Southeast Asia. It plays a significant role in local ecosystems as a food source for various wildlife species and is particularly noted for its traditional medicinal applications in treating different ailments, including digestive disorders and skin conditions (Nguyen et al. 2022; Nguyen, Thi, et al. 2023; Nguyen, Bui, et al. 2024).

Chloroplasts (cps) are essential organelles in plants which house a multitude of proteins crucial for photosynthesis and various metabolic processes (Keeling 2004). The cp genome typically exhibits a quadripartite structure comprising a large single-copy (LSC) region, a small single-copy (SSC) region, and two inverted repeat (IR) regions (Wicke et al. 2011). In recent years, phylogenetic analyses using protein-coding sequences from cp genomes have provided novel insights and significant advancements in our understanding of plant evolution (Asaf et al. 2020; Nhat Nam et al. 2023; Nguyen, Do, et al. 2024). Moreover, the noncoding regions of cp genomes are valuable for plant species identification (Yousefzadeh et al. 2014; Hernández-Ledesma and Bárcenas 2017; Ngai et al. 2023; Wang et al. 2024). Despite the importance of *S. zeylanicum*, there is a notable lack of cp genomic data for this species, which hinders our understanding of its genetic makeup and evolutionary relationships within the *Syzygium* genus. Therefore, we aimed to sequence, assemble, and annotate the complete cp genome of *S. zeylanicum*. This study provides essential genomic information and contributes to resolving potential taxonomic ambiguities within the *Syzygium* genus. Furthermore, by leveraging the cp genome data for phylogenetic analysis, our research provides reliable genetic support for species identification, molecular marker development, and systematic evolutionary studies within *Syzygium* and the broader Myrtaceae family.

Materials

Fresh leaves of *S. zeylanicum* were collected in Tay Ninh Province, Vietnam (11°26′34.2″N 106°12′25.3″E) (Figure 1). Healthy, mature leaves were selected, rinsed with distilled water, dried with silica-gel, and stored at room temperature. A voucher specimen (UMP_2024.04.01_Tramvodo) was deposited at the University of Medicine and Pharmacy at Ho Chi Minh City (contact person: Quang Trong Minh, email: qtminh@ump.edu.vn). Total genomic DNA was extracted from the dried leaf tissue using a modified cetyltrimethylammonium bromide (CTAB) method (Porebski et al. 1997). The extracted DNA was further purified using a Monarch Genomic DNA Purification Kit (#T3010, New England Biolabs,

CONTACT Hoang Danh Nguyen nhdanh@ntt.edu.vn Functional Genomics Research Center, NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City 70000, Vietnam

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2435905.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

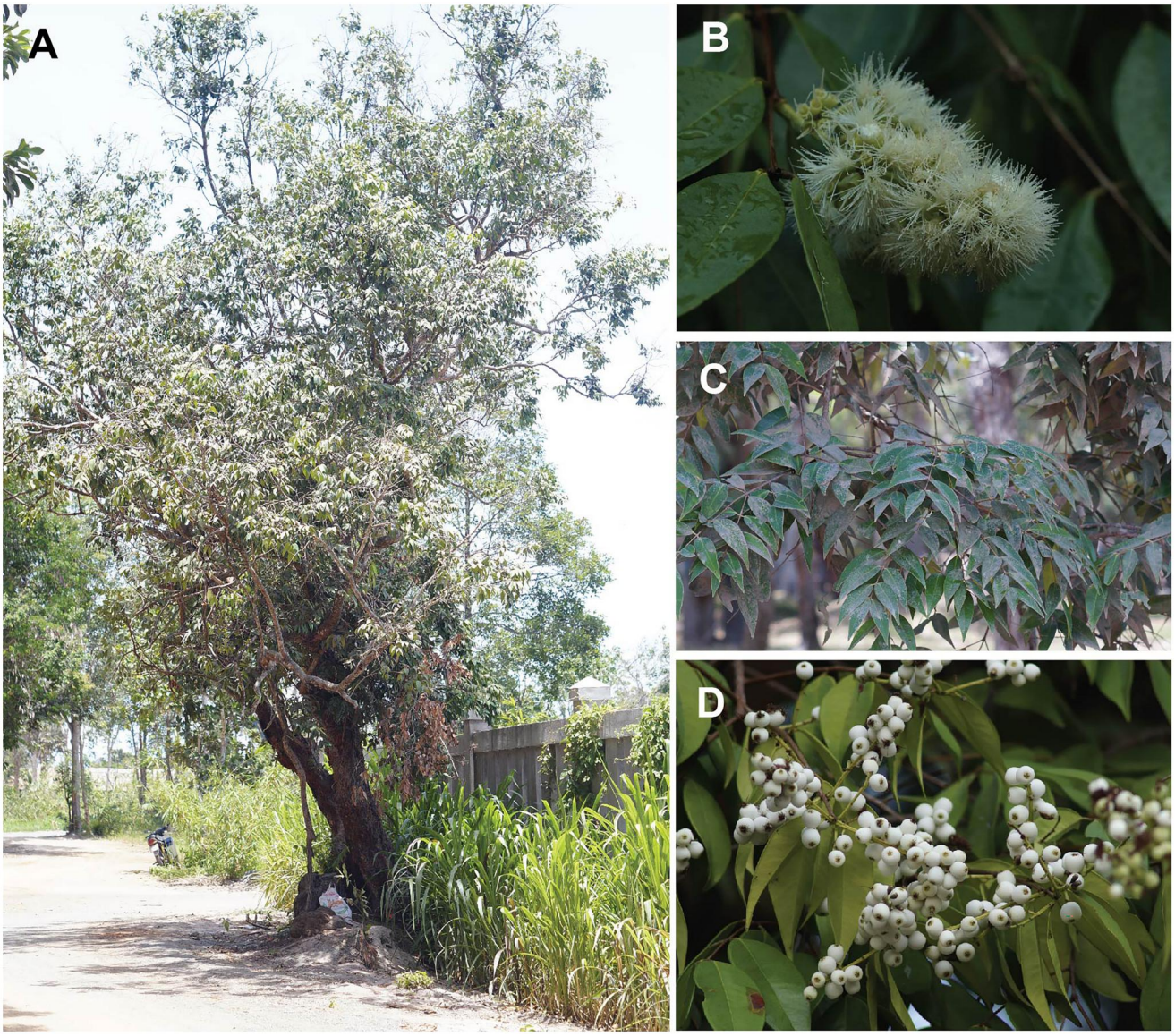

Figure 1. Photographs of *Syzygium zeylanicum*. (A) Plant habitat and entire plant – woody stem, 5–10 m tall, upright, with a round cross-section. Young stems are green and smooth, while older stems become reddish-brown. (B) Inflorescence – cymose inflorescence, growing in leaf axils and at branch tips. Flowers are regular, bisexual, with four sepals. The peduncle is cylindrical, light green, and smooth. There are four petals that are fused at the tip, forming a greenish-white point. Stamens are numerous, separate, and arranged in multiple whorls around the mouth of the flower base. (C) Leaves – simple, opposite, and lack stipules. The leaf blade is thick, hard, and oblong-elliptic with a pointed tip. It is green, darker on the upper surface than the lower, with an entire margin. The leaf veins are pinnate, with numerous lateral veins connecting near the leaf edge. The petiole is cylindrical, slightly twisted, with shallow grooves on the upper surface. (D) Fruits – fleshy berries, white, oval-shaped (photo taken by Minh Trong Quang).

Ipswich, MA) following the manufacturer's protocol. DNA quality and quantity were assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and Qubit fluorometer (Invitrogen, Carlsbad, CA), respectively. High-quality DNA (A260/280 ratio of 1.8–2.0) approximately 50 ng/μL concentration was used for subsequent library preparation.

Methods

Sequencing, assembly, and annotation of the cp genome

The sequencing library was prepared using the NEBNext Ultra II DNA Library Prep kit (#E7645, New England Biolabs, Ipswich, MA). Paired-end sequencing (150 bp in length) was performed on a MiSeq sequencer (Illumina, San Diego, CA). Then, quality control was performed using FastQC v0.11.9 (Brown et al. 2017) with default parameters, followed by adapter trimming and quality filtering using Trimmomatic v0.39 (Bolger et al. 2014). The cp genome was *de novo* assembled using GetOrganelle v1.7.7.0 (Jin et al. 2020) and NOVOPlasty v4.3.1 (Dierckxsens et al. 2017), with *Syzygium polyanthum* (Wight) Walp. 1843 (GenBank accession no. NC_072979) as the reference. Annotation was conducted using GeSeq (Tillich et al. 2017), followed by manual curation and verification of the protein-coding genes (PCGs) and tRNA genes using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and tRNAscan-SE v2.0, respectively (Chan et al. 2021). A circular cp genome map was generated using OGDRAW v1.3.1 (Greiner et al. 2019); cis-splicing and trans-splicing gene maps were constructed using

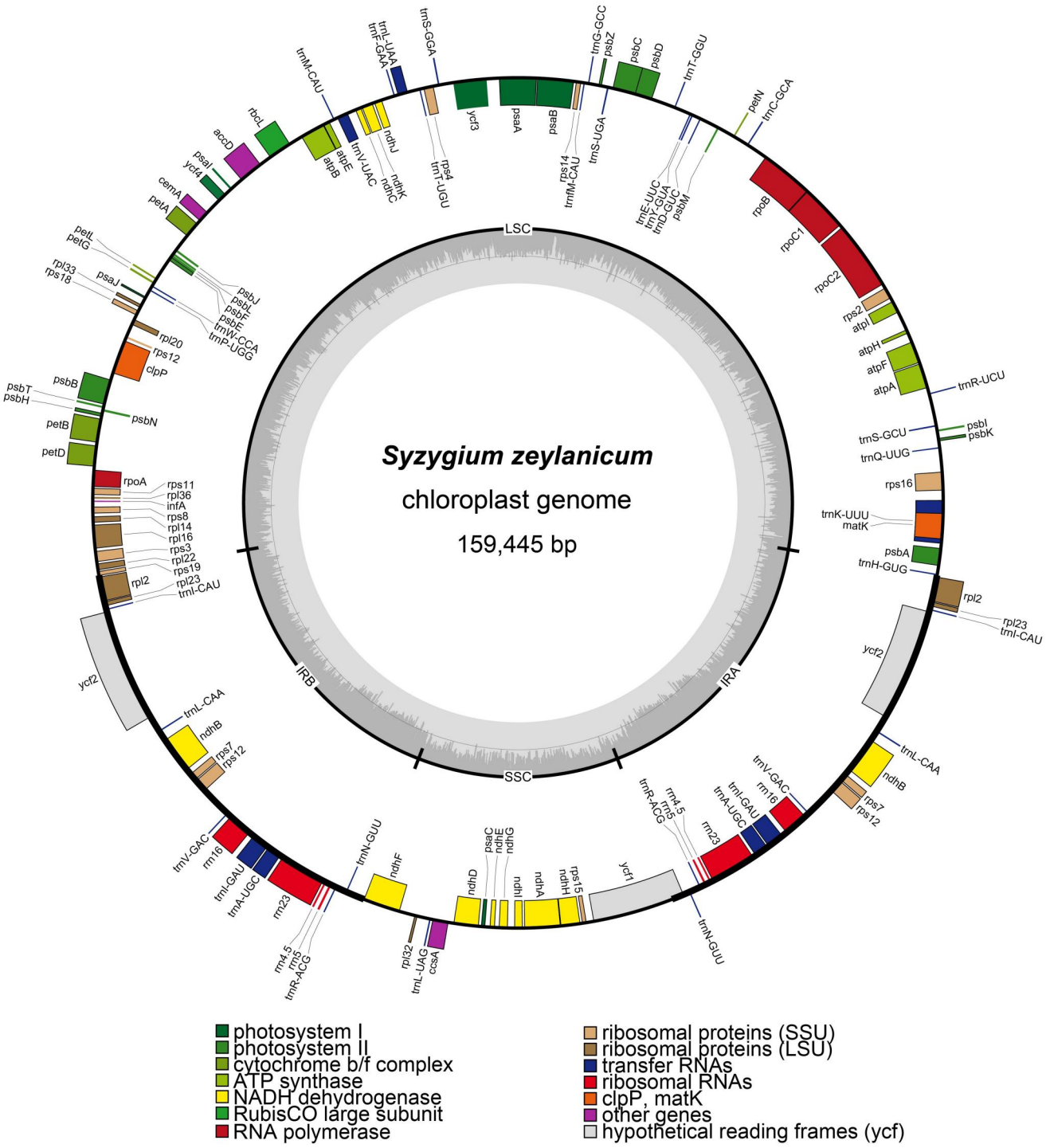

Figure 2. Chloroplast genome map of *Syzygium zeylanicum*. The cp genome consists of a typical four-region circular molecule with a large single-copy (LSC) region, a small single-copy (SSC) region, and a pair of inverted repeats (IRa and IRb) regions. Genes on the outside of the circle are transcribed in a clockwise direction, whereas genes on the inside are transcribed in a counterclockwise direction. The inner circle represents the GC (dark gray) and AT content (light gray). The functional classification of genes is indicated by color coding, as shown in the two columns.

CPGview (Liu et al. 2023). The annotated cp genome was submitted to GenBank (GenBank accession no. PP866890).

Phylogenetic analysis

For the phylogenetic analysis, 13 complete cp genomes of *Syzygium* species were retrieved from the GenBank database and *Stockwellia quadrifida* D.J.Carr, S.G.M.Carr & B.Hyland 2002 (accession no. NC_022414) was selected as an outgroup (Bayly et al. 2013). The PCGs of all sampled species were extracted and aligned using MAFFT v7.490 with the default parameters (Katoh and Standley 2013). Gaps and poor alignment regions were removed using TrimAl v1.2 (Capella-Gutiérrez et al. 2009). The best-fit nucleotide substitution model was determined using ModelFinder, which was implemented in IQ-TREE v2.2.2.6 (Kalyaanamoorthy et al. 2017). Phylogenetic reconstruction was performed using the

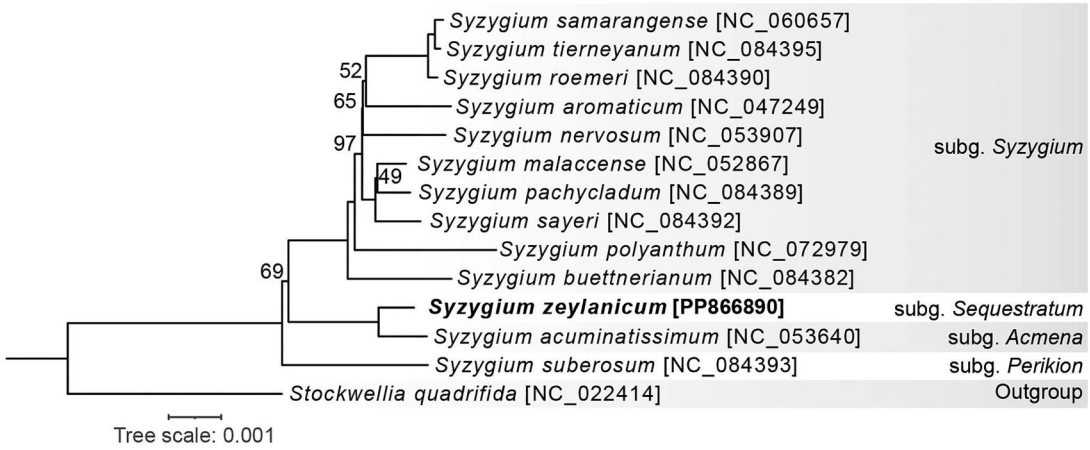

Figure 3. Maximum-likelihood phylogenetic tree of the *Syzygium* genus based on concatenated chloroplast protein-coding genes. The bootstrap values of 100 are omitted. The names of subgenera are depicted at the right of corresponding species. The sequences used in the phylogenetic analysis include *Syzygium samarangense* (NC_060657) (Wei et al. 2022), *Syzygium tierneyanum* (NC_084395), *Syzygium roemeri* (NC_084390), *Syzygium pachycladum* (NC_084389), *Syzygium sayeri* (NC_084392), *Syzygium malaccense* (NC_052867) (Tao et al. 2020), *Syzygium aromaticum* (NC_047249), *Syzygium nervosum* (NC_053907) (Liu et al. 2024), *Syzygium polyanthum* (NC_072979) (Nguyen, Vu, et al. 2023), *Syzygium buettnerianum* (NC_084382), *Syzygium acuminatissimum* (NC_053640) (Zeng et al. 2021), *Syzygium suberosum* (NC_084393), and *Stockwellia quadrifida* (NC_022414) (Bayly et al. 2013).

maximum-likelihood method in IQ-TREE, specifying the molecular substitution models GTR + I + G, with the bootstrap set to 1000 replicates (Nguyen et al. 2015). The resulting phylogenetic tree was visualized and edited using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Results

The *S. zeylanicum* cp genome was successfully assembled and annotated, resulting in a circular genome of 159,445 bp with an average coverage depth of 2352× (Figure S1). The genome exhibited a typical quadripartite structure, consisting of a LSC region of 88,034 bp in length, an SSC region of 18,455 bp in length, and two IR regions (IRa and IRb) of 26,478 bp in length (Figure 2). The overall GC content of the *S. zeylanicum* cp genome was 36.9%, with regional variations as follows: 34.8% in the LSC, 30.7% in the SSC, and 42.7% in the IR regions. Coding regions (comprising PCGs, tRNAs, and rRNAs) constitute 56.8% of the genome. Conversely, noncoding regions, including introns and intergenic spacers, account for 43.2% of the genomic content.

The *S. zeylanicum* cp genomes encoded 130 genes, including 85 PCGs, eight rRNAs, and 37 tRNAs. Of these, 17 genes were found to contain introns, with three genes (*ycf3* and *clpP*) containing two introns and 15 genes having a single intron (Figure S2). Additionally, 17 genes were duplicated in the IR regions, including six PCGs (*ndhB*, *rpl2*, *rpl23*, *rps7*, *rps12*, and *ycf2*), seven tRNA genes (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*), and four rRNA genes (*rrn5*, *rrn4.5*, *rrn23*, and *rrn16*). Notably, the *rps12* gene underwent trans-splicing, a characteristic post-transcriptional modification process in cp genomes (Figure S3).

Phylogenetic analysis, employing the maximum-likelihood method and incorporating data from other *Syzygium* species, was conducted to determine the relationships of *S. zeylanicum* within the genus. The results demonstrated that *S. zeylanicum* and *S. acuminatissimum* (Blume) DC. 1828 formed a distinct clade with strong supported value (bootstrap of 100%) (Figure 3). The robustness of this topology within the genus *Syzygium* is supported by high bootstrap values at each branch node, providing strong evidence for the evolutionary relationships among these species.

Discussion and conclusions

In this study, we successfully sequenced, assembled, and annotated the cp genome of *S. zeylanicum*, providing the first comprehensive characterization of its genomic features. The *S. zeylanicum* cp genome exhibits a typical quadripartite structure with a total length of 159,445 bp, which aligns closely with the range observed in other *Syzygium* species (Wei et al. 2022; Nguyen, Vu, et al. 2023). Furthermore, phylogenetic analysis demonstrated that *S. zeylanicum* was closely related to *S. acuminatissimum* with strong support values (bootstrap value of 100). The utilization of the complete cp genome in these analysis yielded higher resolution compared to traditional approaches employing a limited number of DNA markers (Tao et al. 2020; Li et al. 2021; Zhang et al. 2023; Liu et al. 2024). In conclusion, our study provides a genetic resource for phylogenetic and evolutionary studies of *Syzygium* genus.

Author contributions

MTQ collected the samples. HDN developed the methodology. TTTH and HDN conducted the experiments. TTTH and MTQ wrote the original draft. HDN and MTQ reviewed and edited the manuscript. All authors have read and agreed to the publication of the final version of the manuscript.

Ethical approval

Syzygium zeylanicum is not a protected species under national or international regulations. Therefore, no permits or approvals were required to collect samples from these plants.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Minh Trong Quang was funded by the Master, PhD Scholarship Program of Vingroup Innovation Foundation, Code [VINIF.2021.ThS.69] and [VINIF.2022.ThS.054].

Data availability statement

The genome sequence data supporting the findings of this study are openly available in GenBank under accession number: PP866890 (*Syzygium zeylanicum*). The associated BioProject, SRA, and Bio-Sample numbers were PRJNA1127846, SRX25039856, and SAMN42021377, respectively.

References

Amir Rawa MS, Mazlan MKN, Ahmad R, Nogawa T, Wahab HA. 2022. Roles of *Syzygium* in anti-cholinesterase, anti-diabetic, anti-inflammatory, and antioxidant: from Alzheimer's perspective. Plants. 11(11): 1476. doi:10.3390/plants11111476.

Asaf S, Khan AL, Khan A, Al-Harrasi A. 2020. Unraveling the chloroplast genomes of two *Prosopis* species to identify its genomic information, comparative analyses and phylogenetic relationship. Int J Mol Sci. 21(9):3280. doi:10.3390/ijms21093280.

Bayly MJ, Rigault P, Spokevicius A, Ladiges PY, Ades PK, Anderson C, Bossinger G, Merchant A, Udovicic F, Woodrow IE, et al. 2013. Chloroplast genome analysis of Australian eucalypts – Eucalyptus, Corymbia, Angophora, Allosyncarpia and Stockwellia (Myrtaceae). Mol Phylogenet Evol. 69(3):704–716. doi:10.1016/J.YMPEV.2013.07.006.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15):2114–2120. doi:10.1093/bioinformatics/btu170.

Brown J, Pirrung M, McCue LA. 2017. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics. 33(19):3137–3139. doi:10.1093/bioinformatics/btx373.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25(15):1972–1973. doi:10.1093/bioinformatics/btp348.

Chan PP, Lin BY, Mak AJ, Lowe TM. 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. 49(16):9077–9096. doi:10.1093/nar/gkab688.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18. doi:10.1093/nar/gkw955.

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

Hernández-Ledesma P, Bárcenas RT. 2017. Phylogenetic utility of the trn H–psb A IGR and stem-loop diversity of the 3′ UTR in Cactaceae (Caryophyllales). Plant Syst Evol. 303(3):299–315. doi:10.1007/s00606-016-1372-9.

Jin JJ, Yu W, Bin Yang JB, Song Y, DePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21(1):241. doi:10.1186/s13059-020-02154-5.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von HA, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14(6):587–589. doi:10.1038/nmeth.4285.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780. doi:10.1093/MOLBEV/MST010.

Keeling PJ. 2004. Diversity and evolutionary history of plastids and their hosts. Am J Bot. 91(10):1481–1493. doi:10.3732/ajb.91.10.1481.

Li P, Guo W, Lei K, Ji L. 2021. Characterization of the complete chloroplast genome of *Syzygium nervosum*. Mitochondrial DNA B Resour. 6(3):1014–1015. doi:10.1080/23802359.2021.1894999.

Liu F, Yuan L, Zhang Y. 2024. The complete chloroplast genome of the *Syzygium buxifolium* Hook. Et Arn. 1833 and its phylogenetic analysis. Mitochondrial DNA B Resour. 9(7):851–855. doi:10.1080/23802359.2024.2371553.

Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, Liu C. 2023. CPGView: a package for visualizing detailed chloroplast genome structures. Mol Ecol Resour. 23(3):694–704. doi:10.1111/1755-0998.13729.

Ngai H-L, Kong BL-H, Lau DT-W, Shaw P-C. 2023. Differentiation of Lingxiaohua and Yangjinhua by chloroplast genome sequencing and DNA barcoding markers. Genome. 66(2):21–33. doi:10.1139/gen-2022-0063.

Nguyen HD, Do HDK, Vu MT. 2024. Comparative genomics revealed new insights into the plastome evolution of Ludwigia (Onagraceae, Myrtales). Sci Prog. 107(3):368504241272741. doi:10.1177/00368504241272741.

Nguyen HD, Vu MT, Do HDK. 2023. The complete chloroplast genome of *Syzygium polyanthum* (Wight) Walp. (Myrtales: Myrtaceae). J Asia-Pacific Biodivers. 16(2):267–271. doi:10.1016/j.japb.2023.03.002.

Nguyen L-T, Schmidt HA, Haeseler A, Von Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274. doi:10.1093/molbev/msu300.

Nguyen M, Thi BHB, Maskey S, Tran M, Nguyen Q. 2023. In vitro and in vivo antioxidant and antihyperglycemic potentials of phenolic fractions of *Syzygium zeylanicum* (L.) DC trunk-bark. Food Sci Nutr. 11(7): 3875–3884. doi:10.1002/fsn3.3373.

Nguyen M-T, Bui T-B-H, Pham V-H, Tran M-D, Nguyen Q-V. 2024. *Syzygium zeylanicum* (L.) DC. polyphenols exhibit anti-diabetic activity by modulation of *ACC1*, *SGLT1*, and *GLP-1* genes and restoration of gut microbiota in overfeeding and high glucose exposure-induced diabetic zebrafish. J Funct Foods. 112:105921. doi:10.1016/j.jff.2023.105921.

Nguyen M-T, Chuyen H, Van Tran MD, Nguyen Q-V. 2022. Microencapsulation of *Syzygium zeylanicum* (L.) DC. extract using spray drying: effects of wall materials on physicochemical characteristics and biological activities of the microcapsules. Food Process Preserv. 46(7):e16647. doi:10.1111/jfpp.16647.

Nhat Nam N, Hoang Danh N, Minh Thiet V, Do HDK. 2023. New insights into the evolution of chloroplast genomes in *Ochna* species (Ochnaceae, Malpighiales). Evol Bioinform. 19:11769343231210756. doi:10.1177/11769343231210.

Nigam V, Nigam R. 2012. Distribution and medicinal properties of *Syzygium* species. Curr Res Pharm Sci. 2(2):73–80.

Porebski S, Bailey LG, Baum BR. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Rep. 15(1):8–15. doi:10.1007/BF02772108.

Tao L, Shi Z-G, Long Q-Y. 2020. Complete chloroplast genome sequence and phylogenetic analysis of *Syzygium malaccense*. Mitochondrial DNA B Resour. 5(3):3549–3550. doi:10.1080/23802359.2020.1829132.

Thornhill AH, Ho SYW, Külheim C, Crisp MD. 2015. Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. Mol Phylogenet Evol. 93:29–43. doi:10.1016/j.ympev.2015.07.007.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11. doi:10.1093/NAR/GKX391.

Uddin ABMN, Hossain F, Reza ASMA, Nasrin MS, Alam AHMK. 2022. Traditional uses, pharmacological activities, and phytochemical constituents of the genus *Syzygium*: a review. Food Sci Nutr. 10(6):1789–1819. doi:10.1002/fsn3.2797.

Wang J, Kan S, Liao X, Zhou J, Tembrock LR, Daniell H, Jin S, Wu Z. 2024. Plant organellar genomes: much done, much more to do. Trends Plant Sci. 29(7):754–769. doi:10.1016/j.tplants.2023.12.014.

Wei X, Li L, Xu L, Zhang X, Zeng L, Xu J. 2022. Complete chloroplast genome sequence of *Syzygium samarangense* (Myrtaceae) and phylogenetic analysis. Mitochondrial DNA B Resour. 7(6):977–979. doi:10.1080/23802359.2022.2080022.

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

Yousefzadeh H, Hosseinzadeh Colagar A, Akbarzadeh F, Tippery NP. 2014. Taxonomic status and genetic differentiation of Hyrcanian Castanea based on noncoding chloroplast DNA sequences data. Tree Genet Genomes. 10(6):1611–1629. doi:10.1007/s11295-014-0783-4.

Zeng F, Deng Y, Liu X, Zhu X, Tan G. 2021. The complete chloroplast genome of *Syzygium acuminatissimum*. Mitochondrial DNA B Resour. 6(1):127–128. doi:10.1080/23802359.2020.1847615.

Zhang M, Yang X, Zhu L, Xi D, Cai H, Yin T, Zhang H. 2023. The complete chloroplast genome of two *Syzygium* (Myrtaceae) species and comparative analysis with other related species. Plant Biotechnol Rep. 17(5):753–765. doi:10.1007/s11816-023-00865-2.