

PLASTOME REPORT

 OPEN ACCESS  Check for updates

The complete chloroplast genome sequence of *Enkianthus perulatus* (Miq.) C.K.Schneid. (Ericaceae)

Watanabe Yoichi

Graduate School of Horticulture, Chiba University, Matsudo, Japan

ABSTRACT

Enkianthus perulatus (Miq.) C.K.Schneid. is a popular ornamental plant in Japan, for which the chloroplast genome has not been previously reported. Its complete chloroplast genome was therefore determined, using a high-throughput sequencing technique. The total length of the *E. perulatus* chloroplast genome was 157,256 bp, comprising a large single copy region of 89,680 bp and a small single copy region of 20,472 bp, separated by a pair of 23,552 bp inverted repeat regions. A phylogenetic analysis, based on the chloroplast genome, indicated that *E. perulatus* is the earliest diverged species within Ericaceae. The chloroplast genome structure of *E. perulatus* is similar to species from families within Ericales, but not other species of Ericaceae. The presented chloroplast genome will be useful for further systematic analyses and planning conservation strategies.

ARTICLE HISTORY

Received 29 November 2024
Accepted 26 March 2025


KEYWORDS

Ornamental shrub;
phylogenetic relationships;
rare species; structural
variation



Figure 1. The (A) inflorescence and (B) autumn leaves of *Enkianthus perulatus* plant sampled for the analysis. The photograph was taken by Watanabe Yoichi at Minami-Asai, Chiba Prefecture, Japan. *Enkianthus perulatus* is characterized by an urceolate corolla with a length-to-width ratio of 1.0–1.3, in contrast to a length-to-width ratio of 1.5–2.0 in its closest relative, *E. serrulatus*.

CONTACT Watanabe Yoichi  ashitaka3776@gmail.com  Graduate School of Horticulture, Chiba University, Matsudo, Chiba 2718510, Japan.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2025.2487065>.

© 2025 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.

Introduction

Enkianthus perulatus (Miq.) C.K.Schneid. (Miquel 1863; Schneider 1912) is a deciduous shrub distributed across temperate East Asia (Liang et al. 2022). The species is a popular ornamental shrub in Japan, with beautiful spring flowers and autumn leaves (Figure 1). Its natural habitat is largely restricted to fragmented serpentine terrain in the southwestern area of the main islands of Japan; it has also been

recorded in eastern parts of mainland China and Taiwan (Liang et al. 2022).

Genomic features of the chloroplast genome are useful for analyzing phylogenetic and population genetic relationships of natural populations and cultivated individuals (Zhou et al. 2024). The structure of the chloroplast genome is known to differ among Ericaceae taxa. For example, the genus *Rhododendron* of the subfamily Ericoideae has a large chloroplast genome with repetitive genes, in addition only *ndhF* is

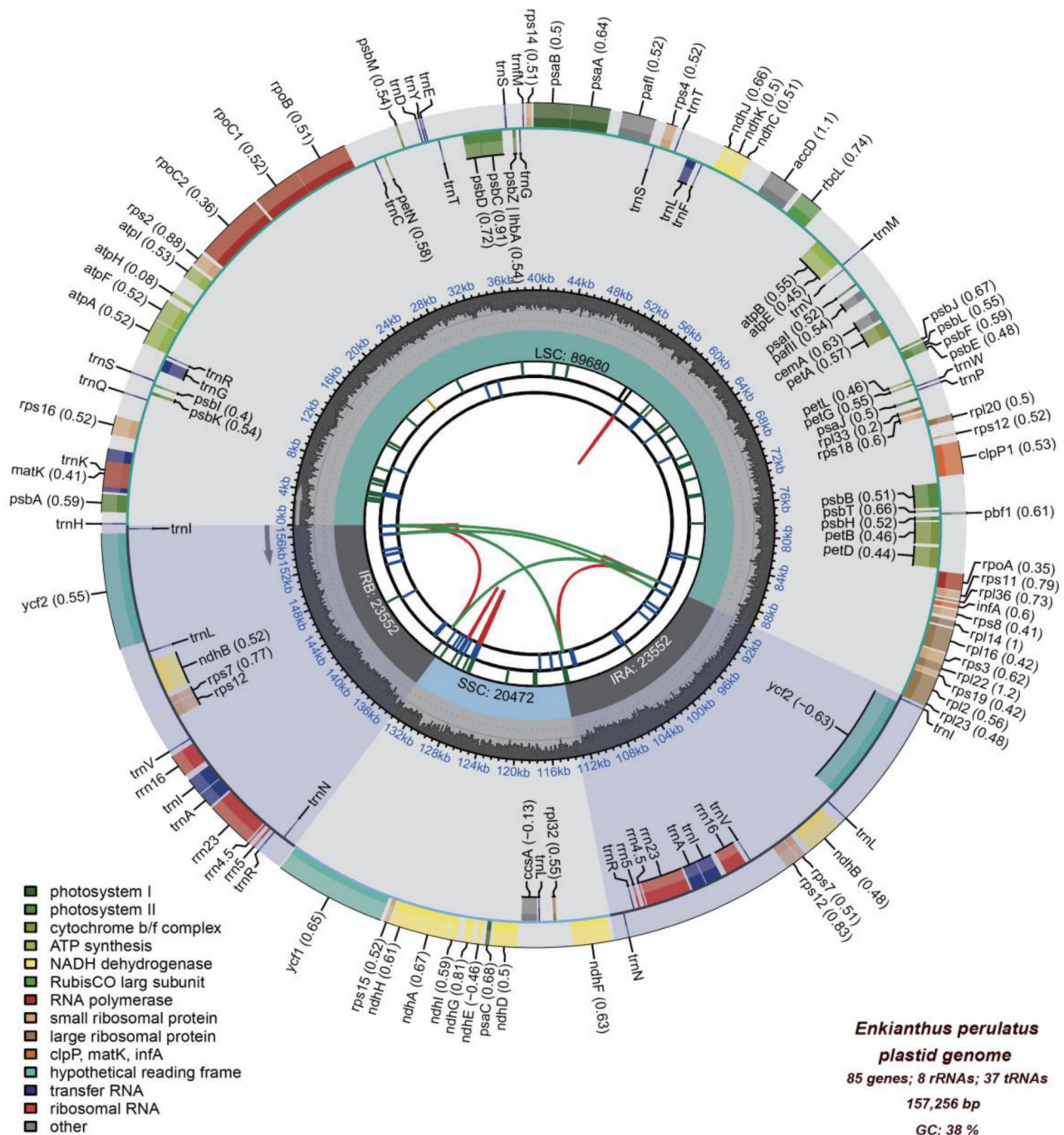


Figure 2. The complete chloroplast genome of *Enkianthus perulatus*. The map contains six tracks. From the center outward, the first track displays the dispersed repeats. The dispersed repeats consist of direct and palindromic repeats, connected with red and green arcs. The second track shows long tandem repeats as short blue bars. The third track shows short tandem repeats or microsatellite sequences as short bars with different colors. The small single-copy (SSC), inverted repeat (IRA and IRB), and large single-copy (LSC) regions are shown on the fourth track. The GC content along the genome is plotted on the fifth track. The genes are shown on the sixth track. Optional codon usage bias is displayed in parentheses after the gene name. Genes belonging to different functional groups are color-coded. Genes on the inside and outside of the map are transcribed in clockwise and counterclockwise directions, respectively.

included in a small single copy (SSC; Zhou et al. 2023). *Enkianthus* is the sole genus of the subfamily Enkianthoideae, which is considered to be the earliest diverged group in Ericaceae (Kron et al. 2002; Schwery et al. 2015). The chloroplast genome feature of the genus is not known; thus, its complete chloroplast genome sequence will provide knowledge for the development of structural variation in Ericaceae. In this study, I report the complete chloroplast genome of *E. perulatus* based on Illumina paired-end sequencing data. In addition, phylogenetic relationships of the genus *Enkianthus* and related genera were reconstructed by utilizing published sequences of related species.

Materials and methods

Sample leaves were collected from Minami-Asai in Chiba Prefecture, Japan (34.9496°N, 139.9312°E) on 14 July 2023 (GenBank BioSample, SAMD00834094). The sample was identified and collected by the author, and the voucher specimen was deposited in the Herbarium of the National Museum of Nature and Science, Japan (<https://tbk.kahaku.go.jp/english/index.php>, Atsushi Ebihara, ebihara@kahaku.go.jp) under the accession number TNS1352511 (Figure 1).

Total genomic DNA was extracted from dried leaves using the DNeasy Plant Mini kit (Qiagen, Hilden, Germany). The extracted DNA was used to prepare a sequencing library using the TruSeq DNA PCR-Free Kit (Illumina, San Diego, CA) and then sequenced with 150 bp of pair-end reads on the HiSeqX platform (Illumina); the library preparation and sequencing were conducted by MacroGen (Seoul, South Korea). The 7.3 Gbases of raw data were trimmed by clipping adaptor sequences and removing reads of low quality using fastp v. 0.20.0 (Chen et al. 2018). Trimmed reads were deposited in GenBank (accession number, DRR620097). The trimmed reads were then assembled using GetOrganelle v.1.7.5 (Jin et al. 2020) with the setting -F embplant_pt, and default settings for other options. The assembled circular sequence was checked using Bandage (Wick et al. 2015). The assembled genome was annotated based on BLAST searches using GeSeq (Tillich et al. 2017) with default settings. To verify complete coverage of the chloroplast genome, the trimmed reads were mapped to the assembled genome using BWA (Li and Durbin 2009) with default settings. To facilitate mapping to the circular genome, the sequence was linearized and 700 bp duplicated from one end to the other. The genome map and cis/trans-splicing genes maps were drawn using CPGView (Liu et al. 2023). The annotated

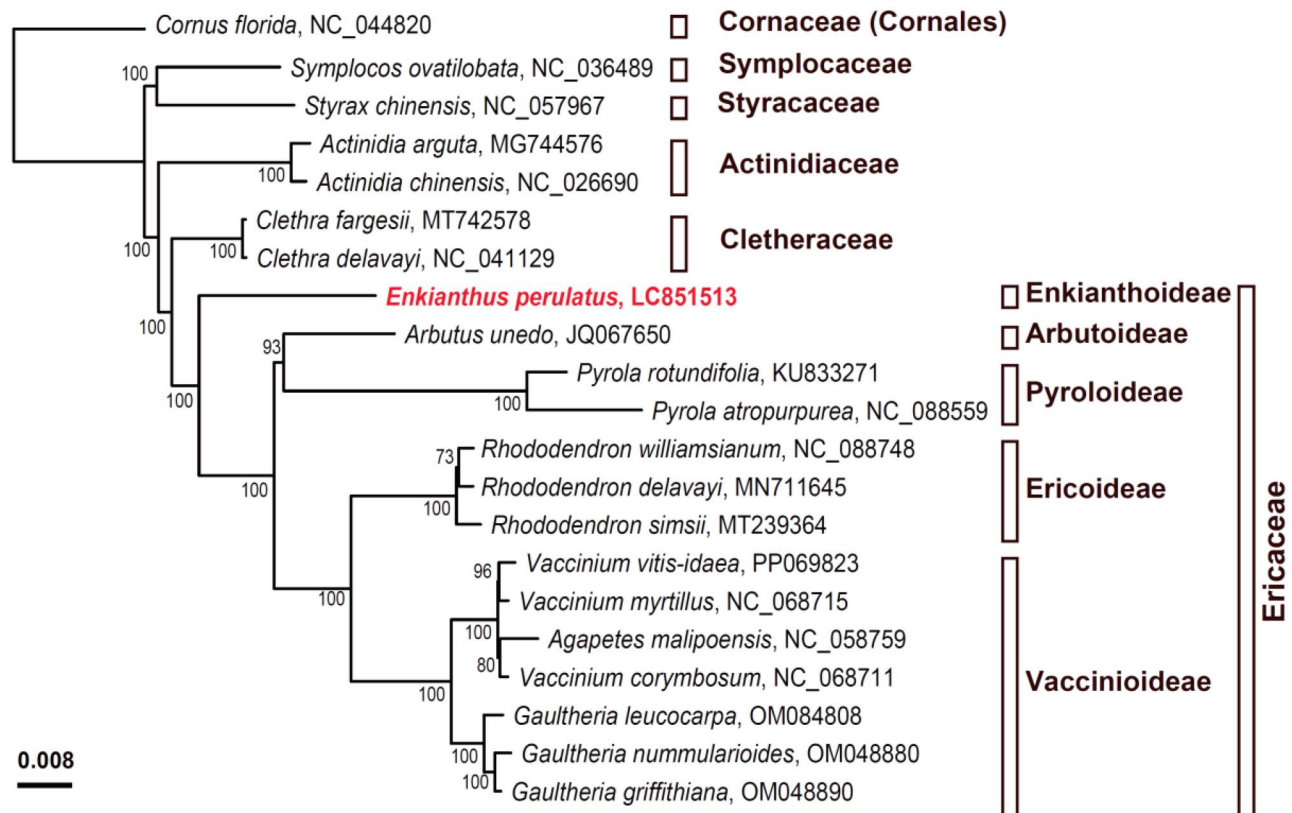


Figure 3. Phylogenetic relationships between *Enkianthus perulatus* and 20 other species from Ericaceae and five other families. Bootstrap probabilities are shown above the nodes. The genome of the species identified with red-bold font is presented in this study. The following sequences were used: *Gaultheria griffithiana* OM048890 (Xu et al. 2022), *G. nummularioides* OM048880 (Xu et al. 2022), *G. leucocarpa* OM084808 (Xu et al. 2022), *Agapetes malipoensis* NC_058759, *Vaccinium corymbosum* NC_068711 (Fahrenkrog et al. 2022), *V. myrtillus* NC_068715 (Fahrenkrog et al. 2022), *V. vitis-idaea* PP069823, *Rhododendron delavayi* MN711645 (Li et al. 2020), *R. simsii* MT239364 (Zhou et al. 2023), *R. williamsianum* NC_088748, *Pyrola atropurpurea* NC_088559, *P. rotundifolia* KU833271 (Logacheva et al. 2016), *Arbutus unedo* JQ067650 (Martínez-Alberola et al. 2013), *Clethra delavayi* NC_041129, *C. fargesii* MT742578 (Ding et al. 2021), *Actinidia chinensis* NC_026690 (Yao et al. 2015), *A. arguta* MG744576 (Lin et al. 2018), *Styrax chinensis* NC_057967 (Tian et al. 2020), *Symplocos ovatilobata* NC_036489 (Zhu et al. 2018), and *Cornus florida* NC_044820 (Fu et al. 2019).

chloroplast genome was also deposited in GenBank (accession number, LC851513).

To investigate the phylogenetic position of *E. perulatus*, chloroplast genome sequences of related taxa were obtained from the GenBank database. Thirteen Ericaceae species from four subfamilies, six species from four families of Ericales, and one species from a family of Cornales were used as outgroups for the phylogenetic analysis. A total of 70 coding

sequences (CDSs) that were shared among the species were extracted, and each orthologous gene was aligned using MAFFT v.7.505 (Kato and Standley 2013) with default settings, and then these alignments combined. The selection of a substitution model and construction of a phylogenetic tree by the maximum likelihood method, with 1000 bootstrap replicates, were made using IQ-TREE web server v.1.6.12 (Trifinopoulos et al. 2016). The synteny across 11 species

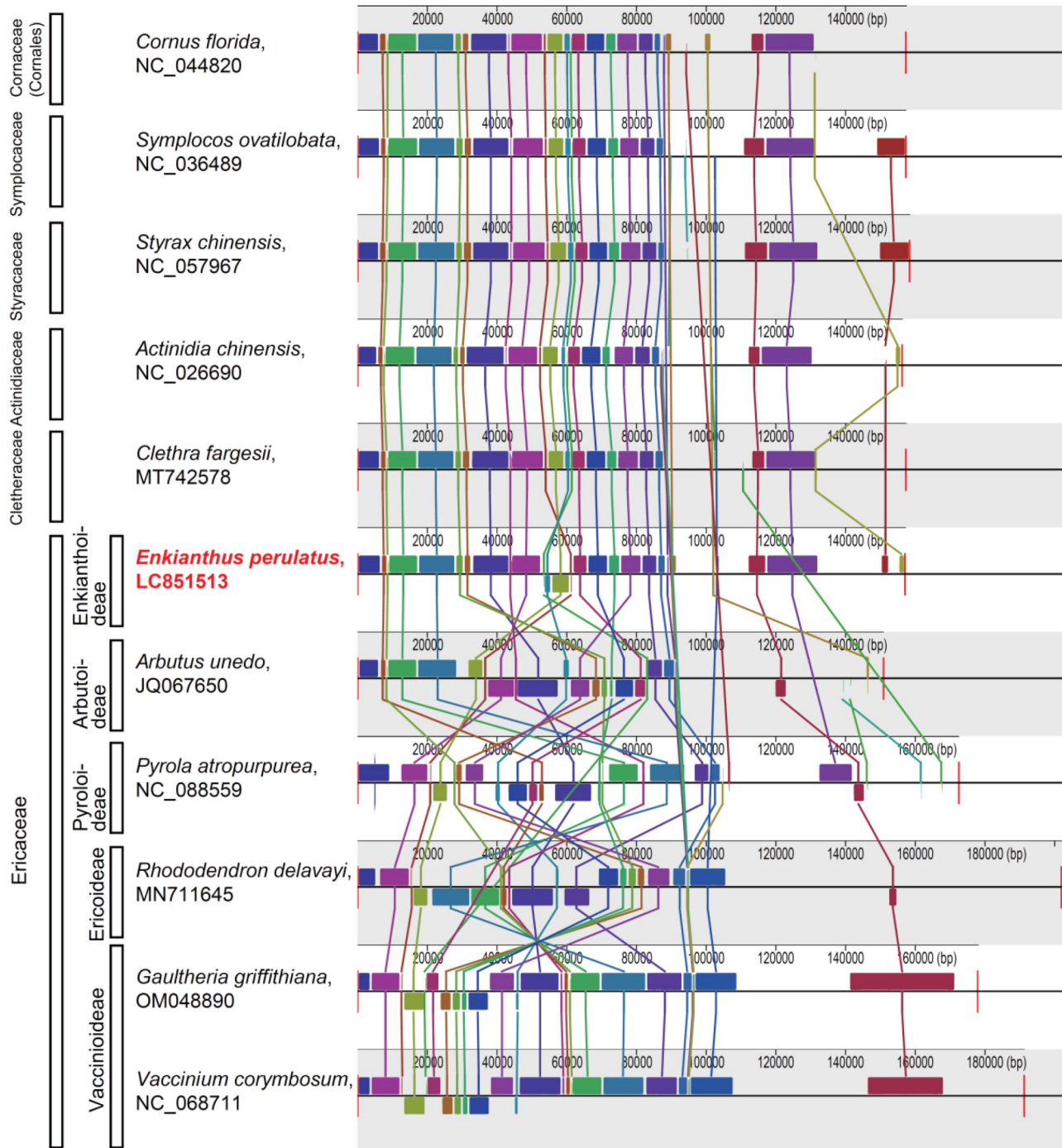


Figure 4. Structural variation in chloroplast genomes across 11 species from the orders Ericales and cornales visualized using mauve. The genome of the species identified with red-bold font was presented in this study. Blocks with the same colors indicate collinear regions, and blocks above and below the center line indicate that they are on the same strand and the opposite strand. The following sequences were used: *Cornus florida* NC_044820 (Fu et al. 2019), *Symplocos ovatilobata* NC_036489 (Zhu et al. 2018), *Styrax chinensis* NC_057967 (Tian et al. 2020), *Actinidia chinensis* NC_026690 (Yao et al. 2015), *Clethra fargesii* MT742578 (Ding et al. 2021), *Arbutus unedo* JQ067650 (Martínez-Alberola et al. 2013), *Pyrola rotundifolia* KU833271 (Logacheva et al. 2016), *Rhododendron delavayi* MN711645 (Li et al. 2020), *Gaultheria griffithiana* OM048890 (Xu et al. 2022), and *Vaccinium corymbosum* NC_068711 (Fahrenkrog et al. 2022).

from Ericales and Cornales was visualized using Mauve v.2.4.0 (Darling et al. 2004). One species from each genus was then chosen from those used in the phylogenetic analysis.

Results

The total length of the complete chloroplast genome of *E. perulatus* was 157,256 bp, with an average coverage of 783.0× and GC content of 38% (Figure 2, Supplementary Figure 1). A large single copy (LSC: 89,680 bp), a small single copy (SSC: 20,472 bp) and two inverted repeat (IR: 23,552 bp) regions made up the quadripartite structure of the chloroplast genome. A total of 114 unique genes were annotated, including 80 protein-coding genes, 30 transfer RNA (tRNA) genes, and four ribosomal RNA (rRNA) genes. The IR regions contained all four rRNA genes, four protein-coding genes and seven tRNA genes, which were therefore duplicated in the chloroplast genomes. Thirteen cis-splicing genes and one trans-splicing gene, *rps12*, were identified (Supplementary Figures 2 and 3).

The phylogenetic relationships between 21 species showed that 14 Ericaceae species were monophyletic with the highest probability (100%, Figure 3). *E. perulatus* was the earliest diverged in Ericaceae. The structural similarity of chloroplast genomes between 11 species showed that *E. perulatus* was highly conserved with species from other families of Ericales and Cornales, in contrast, largely differed from the other Ericaceae species (Figure 4).

Discussion and conclusion

The results of the phylogenetic analysis in this study concur with the topology of Ericaceae presented in previous studies (Kron et al. 2002; Schwery et al. 2015), and indicate that the genus *Enkianthus* is the earliest diverged group in Ericaceae (Figure 3).

This study is the first report of the complete chloroplast genome for the genus *Enkianthus*. The chloroplast genome structure of *E. perulatus* differs from that of other subfamilies in Ericaceae (Figure 4). In detail, the SSC of *E. perulatus* included thirteen genes (Figure 2), which is similar to other families of Ericales (Yao et al. 2015; Zhu et al. 2018; Ding et al. 2021). In contrast, the SSC of *Rhododendron* in subfamily Ericoideae (Zhou et al. 2023) and *Vaccinium* in subfamily Vaccinioideae (Fahrenkrog et al. 2022) only included *ndhF*, while that of *Arbutus* in subfamily Arbutoideae includes *ndhF* and *rpl32* (Martínez-Alberola et al. 2013). In addition, large structural variations of LSC were found within Ericaceae (Figure 4), indicating that such large structural variations were not formed at the establishment of Ericaceae. Although species from all nine subfamilies were not used in the data analysis, previous studies have shown subfamilies Arbutoideae and Pyroloideae to be the second-earliest diverged group in Ericaceae (Kron et al. 2002; Schwery et al. 2015); therefore, the large structural variations may have developed since the divergence of subfamily Enkianthoideae from the others. The size of the chloroplast genome of the mycoheterotrophic plant in the subfamily Monotropoideae

decreases to 35,336 bp (Ravin et al. 2016), and the genomic basis of the large structural variations may have contributed to the ecological diversity of this species-rich family, with an estimated 4,426 species in c. 129 genera of nine subfamilies (Schwery et al. 2015).

This study provides a reliable chloroplast genome for *E. perulatus*. The characteristics of the sequences and genome structure presented can contribute to taxonomic, horticultural and evolutionary studies that focus on *Enkianthus* and Ericaceae. As only the chloroplast genome of the genus has been reported, further studies will help reinforce these findings.

Acknowledgments

W. Y. designed the study, collected the sample, conducted molecular experiments, generated, analyzed, and visualized data, as well as wrote the manuscript.

Author contributions

CRedit: **Watanabe Yoichi**: Conceptualization, Data curation, Formal analysis, Investigation.

Ethical approval

The material involved in this article does not involve any ethical conflicts. This species is not endangered according to the CITES catalog or IUCN Red List, and the sample was not collected from a natural reserve, so the collection did not require any specific permissions or licenses of Japanese governments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by a Grant-in-Aid for Scientific Research for the Japan Society for the Promotion of Science, KAKENHI [Grant No. 20H03020].

Data availability statement

The data that support the findings presented in this study are openly available in GenBank (NCBI, <https://www.ncbi.nlm.nih.gov>) under the reference number LC851513. The associated BioProject, SRA and BioSample numbers are PRJDB19526, DRR620097 and SAMD00834094, respectively.

References

- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34(17):i884–i890. doi:10.1093/bioinformatics/bty560.
- Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. 14(7):1394–1403. doi:10.1101/gr.2289704.
- Ding S, Dong X, Yang J, Guo C, Cao B, Guo Y, Hu G. 2021. Complete chloroplast genome of *Clethra fargesii* Franch., an original sympetalous plant from central China: comparative analysis, adaptive evolution, and phylogenetic relationships. *Forests*. 12(4):441. doi:10.3390/f12040441.
- Fahrenkrog AM, Matsumoto GO, Toth K, Jokipii-Lukkari S, Salo HM, Häggman H, Benevenuto J, Munoz PR. 2022. Chloroplast genome assemblies and comparative analyses of commercially important

- Vaccinium* berry crops. *Sci Rep.* 12(1):21600. doi:[10.1038/s41598-022-25434-5](https://doi.org/10.1038/s41598-022-25434-5).
- Fu CN, Mo ZQ, Yang JB, Ge XJ, Li DZ, Xiang QY, Gao LM. 2019. Plastid phylogenomics and biogeographic analysis support a trans-Tethyan origin and rapid early radiation of Cornales in the Mid-Cretaceous. *Mol Phylogenet Evol.* 140:106601. doi:[10.1016/j.ympev.2019.106601](https://doi.org/10.1016/j.ympev.2019.106601).
- Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21(1):241. doi:[10.1186/s13059-020-02154-5](https://doi.org/10.1186/s13059-020-02154-5).
- 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](https://doi.org/10.1093/molbev/mst010).
- Kron KA, Judd WS, Stevens PF, Crayn DM, Anderberg AA, Gadek PA, Quinn CJ, Luteyn JL. 2002. Phylogenetic classification of Ericaceae: molecular and morphological evidence. *Bot Rev.* 68:335–423. doi:[10.1663/0006-8101\(2002\)068\[0335:PCOEMA\]2.0.CO;2](https://doi.org/10.1663/0006-8101(2002)068[0335:PCOEMA]2.0.CO;2).
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 25(14):1754–1760. doi:[10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324).
- Li H, Guo Q, Li Q, Yang L. 2020. Long-reads reveal that *Rhododendron delavayi* plastid genome contains extensive repeat sequences, and recombination exists among plastid genomes of photosynthetic Ericaceae. *PeerJ.* 8:e9048. doi:[10.7717/peerj.9048](https://doi.org/10.7717/peerj.9048).
- Liang H, Jiang L, Li D, Yang Y, Fan D, Zhang Z. 2022. A new synonym of *Enkianthus perulatus* (Ericaceae) in East Asia, based on morphological and molecular evidence. *PhytoKeys.* 214:61–74. doi:[10.3897/phytokeys.214.94294](https://doi.org/10.3897/phytokeys.214.94294).
- Lin M, Qi X, Chen J, Sun L, Zhong Y, Fang J, Hu C. 2018. The complete chloroplast genome sequence of *Actinidia arguta* using the PacBio RS II platform. *PLoS One.* 13(5):e0197393. doi:[10.1371/journal.pone.0197393](https://doi.org/10.1371/journal.pone.0197393).
- 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](https://doi.org/10.1111/1755-0998.13729).
- Logacheva MD, Schelkunov MI, Shtratnikova VY, Matveeva MV, Penin AA. 2016. Comparative analysis of plastid genomes of non-photosynthetic Ericaceae and their photosynthetic relatives. *Sci Rep.* 6(1):30042. doi:[10.1038/srep30042](https://doi.org/10.1038/srep30042).
- Martínez-Alberola F, del Campo EM, Lázaro-Gimeno D, Mezquita-Claramonte S, Molins A, Mateu-Andrés I, Pedrola-Monfort J, Casano LM, Barreno E. 2013. Balanced gene losses, duplications and intensive rearrangements led to an unusual regularly sized genome in *Arbutus unedo* chloroplasts. *PLoS One.* 8(11):e79685. doi:[10.1371/journal.pone.0079685](https://doi.org/10.1371/journal.pone.0079685).
- Miquel FAG. 1863. *Andromeda perulata*. *Ann Musei Bot Lugduno-Batavi.* 1:31. <https://www.biodiversitylibrary.org/bibliography/113>.
- Ravin NV, Gruzdev EV, Beletsky AV, Mazur AM, Prokhortchouk EB, Filyushin MA, Kochieva EZ, Kadnikov VV, Mardanov AV, Skryabin KG. 2016. The loss of photosynthetic pathways in the plastid and nuclear genomes of the nonphotosynthetic mycoheterotrophic eudicot *Monotropa hypopitys*. *BMC Plant Biol.* 16(Suppl 3):238. doi:[10.1186/s12870-016-0929-7](https://doi.org/10.1186/s12870-016-0929-7).
- Schneider CK. 1912. *Enkianthus* in *Illustriertes Handbuch der Laubholzkunde: Charakteristik der in Mitteleuropa heimischen und im freien angepflanzten angiospermen Gehölz-Arten und Formen mit Ausschluss der Bambuseen und Kakteen*, Band II, pp.520–521, Verlag von Gustav Fischer, Jena, Germany. <https://www.biodiversitylibrary.org/bibliography/194>.
- Schwery O, Onstein RE, Bouchenak-Khelladi Y, Xing Y, Carter RJ, Linder HP. 2015. As old as the mountains: the radiations of the Ericaceae. *New Phytol.* 207(2):355–367. doi:[10.1111/nph.13234](https://doi.org/10.1111/nph.13234).
- Tian Y, Zhang Y, Tong L, Xu X, Wang Y, Jiang X, Wang H. 2020. The complete chloroplast genome sequence of *Styrax chinensis* Hu et S.Y. Liang (Styracaceae). *Mitochondrial DNA B Resour.* 5(3):3363–3365. doi:[10.1080/23802359.2020.1821822](https://doi.org/10.1080/23802359.2020.1821822).
- 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](https://doi.org/10.1093/nar/gkx391).
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44(W1):W232–W235. doi:[10.1093/nar/gkw256](https://doi.org/10.1093/nar/gkw256).
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualisation of de novo genome assemblies. *Bioinformatics.* 31(20):3350–3352. doi:[10.1093/bioinformatics/btv383](https://doi.org/10.1093/bioinformatics/btv383).
- Xu YL, Shen HH, Du XY, Lu L. 2022. Plastome characteristics and species identification of Chinese medicinal wintergreens (*Gaultheria*, Ericaceae). *Plant Divers.* 44(6):519–529. doi:[10.1016/j.pld.2022.06.002](https://doi.org/10.1016/j.pld.2022.06.002).
- 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(6):e0129347. doi:[10.1371/journal.pone.0129347](https://doi.org/10.1371/journal.pone.0129347).
- Zhong S. 2018. *Enkianthus perulatus* in illustrated flora of Taiwan, Volume 6: cornaceae–Bignoniaceae. Taipei, Taiwan: Maotouying Press; p. 45.
- Zhou X, Liu M, Song L. 2023. Structural characterization of four *Rhododendron* spp. Chloroplast genomes and comparative analyses with other azaleas. *Biocell.* 47(3):657–668. doi:[10.32604/biocell.2023.026781](https://doi.org/10.32604/biocell.2023.026781).
- Zhou C, Liang H, Hu W, Yang Y, Tian X, Tsutsumi C, Fan D, Zhang Z. 2024. Molecular phylogeny of *Enkianthus* Lour. (Ericaceae) based on chloroplast and nuclear DNA sequences with an emphasis on the origin of polyploid species. *Sci Hortic.* 328:112960. doi:[10.1016/j.scienta.2024.112960](https://doi.org/10.1016/j.scienta.2024.112960).
- Zhu ZX, Wang JH, Cai YC, Zhao KK, Zhou RC, Wang HF. 2018. Characterization of the complete chloroplast genome sequence of *Symplocos ovatilobata* (Symplocaceae). *Conservation Genet Resour.* 10(3):503–506. doi:[10.1007/s12686-017-0860-8](https://doi.org/10.1007/s12686-017-0860-8).