

First chloroplast genome of *Scleropyrum wallichianum* (Cervantesiaceae) insights into phylogeny and conservation

THUY LIEN BUI¹, LINYUE HE², WYCKLIFFE OMONDI OMOLLO³, SHENGDAN WU², BA CUONG CAO¹, XUAN THANH NGUYEN¹, XUAN THANH TRINH⁴, PHI BANG CAO⁵, NGAN THI LU⁶, CHI TOAN LE¹✉

¹Hanoi Pedagogical University 2, Nguyen Van Linh Street, Xuan Hoa 280000, Phu Tho Province, Vietnam. Tel./fax.: +84-977-522153,

✉email: Lechitoan@hpu2.edu.vn

²State Key Laboratory of Herbage Improvement and Grassland Agro-ecosystems, College of Ecology, Lanzhou University, 222 Tianshui South Rd., Lanzhou 730000, China

³National Key Laboratory for Germplasm Innovation and Utilization of Horticultural Crops, Huazhong Agricultural University, No. 1 Shizishan St., Wuhan 430070, China

⁴Institute of Biology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet St., Nghia Do, Cau Giay, Hanoi, Vietnam

⁵Faculty of Natural Sciences, Hung Vuong University, Nguyen Tat Thanh St., Viet Tri 35120, Phu Tho, Vietnam

⁶Vietnam National Museum of Nature, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet St., Cau Giay, Hanoi, Vietnam

Manuscript received: 1 July 2025. Revision accepted: 10 November 2025.

Abstract. Bui TL, He L, Omollo WO, Wu S, Cao BC, Nguyen XT, Trinh XT, Cao PB, Lu NT, Le CT. 2025. First chloroplast genome of *Scleropyrum wallichianum* (Cervantesiaceae) insights into phylogeny and conservation. *Biodiversitas* 26: 5763-5769. *Scleropyrum wallichianum*, a member of the family Cervantesiaceae (order Santalales), is distributed across Southeast Asia and China. The species is economically significant due to its seed oil, which is used in the production of lubricants and soap. In this study, the complete chloroplast genome of *S. wallichianum* was sequenced and analyzed to clarify its phylogenetic position within Santalales and to support conservation efforts. A leaf sample was collected from the Me Linh Station for Biodiversity in northern Vietnam. The genome was sequenced using the DNBSEQ-T7 platform, assembled with GetOrganelle, and annotated with CPGAVAS2, with manual correction of ambiguous regions performed in Apollo. The circular plastome was visualized using OrganellarGenomeDRAW. The assembled chloroplast genome, representing the first complete sequence reported for *S. wallichianum*, provides a valuable genomic resource for molecular systematics, evolutionary genomics, and conservation genetics of Cervantesiaceae and the order Santalales. The plastome exhibits a typical quadripartite structure with a total length of 149,066 bp and a GC content of 38.3%, comprising a Large Single-Copy (LSC) region of 85,000 bp, a Small Single-Copy (SSC) region of 18,500 bp, and two Inverted Repeats (IRs) of 22,783 bp each. Phylogenomic reconstruction based on 80 shared coding genes strongly supports *S. wallichianum* as a distinct clade within Cervantesiaceae, sister to Opiliaceae. This newly characterized plastome enriches the genomic resources available for the genus *Scleropyrum* and enhances understanding of the evolutionary history, taxonomy, and conservation genetics of Cervantesiaceae and related parasitic lineages within Santalales.

Keywords: Cervantesiaceae, chloroplast, genome, molecular phylogeny, *Scleropyrum wallichianum*

INTRODUCTION

Biodiversity conservation seeks to preserve species, genetic diversity, and ecosystems in order to maintain ecological stability and sustain essential ecosystem services in the face of accelerating environmental change (Cardinale et al. 2012; Pimm et al. 2014). As global biodiversity declines due to habitat degradation, overexploitation, and climate shifts, conservation strategies increasingly rely on molecular and genomic data to clarify species' evolutionary distinctiveness, and demographic histories that support evidence-based management (Allendorf et al. 2010). Chloroplast genomic data play a critical role in this context because plastomes provide conserved yet phylogenetically informative markers that can be used to assess genetic variation, population connectivity, and historical biogeography—key factors in identifying conservation units and protecting declining populations (Daniell et al. 2016; Li et al. 2019). Moreover, chloroplast genomes support species authentication, monitoring of trade, and restoration

provenance analysis, while also improving phylogenetic resolution for lineages where morphological traits are limited or convergent, thereby strengthening conservation prioritization and long-term management planning (Hollingsworth et al. 2011; Coissac et al. 2016).

Hence, the chloroplast genome is a key resource for elucidating phylogenetic relationships among plant species, significantly advancing our understanding of plant evolution (Peruzzi 2016) as well as biodiversity conservation management. Because of its relatively conserved gene content and structure, along with uniparental inheritance, the plastome is especially valuable for resolving complex or unclear nodes in angiosperm phylogenies, including relationships at the base of the phylogeny (Choi et al. 2024). The increasing availability of complete plastome sequences has greatly facilitated phylogenomic studies across diverse plant lineages, thereby deepening our insights into evolutionary patterns at both broad and fine taxonomic levels.

The family Cervantesiaceae (Santalales), comprising 8 genera including *Scleropyrum*, exhibits unique morphological and genetic traits that are crucial for understanding their evolutionary dynamics and classification within Santalales. *Scleropyrum wallichianum*, one of the six recognized species in the genus *Scleropyrum*, exhibits distinctive morphological characteristics, such as variation in spine presence, which are significant for taxonomic classification and evolutionary studies (Figure 1). Xia and Gilbert (2003) noted that the oil extracted from the seeds of *S. wallichianum* can be used in the production of lubricants and soap. *S. wallichianum* is an economically and ecologically important species distributed across tropical and subtropical Asia (Le et al. 2024), yet genomic data for this lineage have been largely unavailable. By obtaining the plastome sequence, researchers could create a key genomic reference that may aid in resolving its systematic placement within the order Santalales, which includes many parasitic and hemiparasitic plants with complex evolutionary histories (Le 2018). While this species has been the subject of several studies (Le et al. 2019, 2024; Zhou et al. 2019), its complete plastome has not yet been characterized.

Guo et al. (2020) studied the plastomes of five species—*Osyris wightiana*, *Pyralaria edulis*, *Santalum album*, *Viscum liquidambaricolum*, and *Viscum ovalifolium* and suggested that reduced genome size, structural rearrangements, and gene loss are characteristic features of hemiparasite plastomes. The loss or pseudogenization of *ndh* genes is commonly observed in Santalales hemiparasites, likely due to the shift in lifestyle from photoautotrophy to hemiparasitism. However, the study also found that the degree of plastome reduction does not correlate with the trophic level, whether facultative or obligate hemiparasitism.

Although Jin et al. (2023) used the plastomes of *S. pentandrum* to investigate diversity patterns of woody angiosperms in China, their analysis did not focus on chloroplast genome structure. In contrast, a complete chloroplast genome analysis indicates that species in the *Scleropyrum* genus exhibit unique genome structures. Given these findings, an in-depth investigation of the chloroplast genome structure of *S. wallichianum* is essential for advancing our understanding of plastome evolution in Cervantesiaceae. *S. wallichianum* is a root-parasitic species, and Nickrent et al. (2010), Nickrent (2011), and Le (2018) have noted that root-parasites represent an intermediate life form, bridging autotrophic and aerial parasitic lifestyles in Santalales. The genetic material of root-parasites contains unique features that are key to understanding the evolution of parasitic plants. Furthermore, Guo et al. (2020) highlighted the rich diversity within Santalales, though they pointed out that only a small fraction of the species in this order have had their plastomes sequenced. Given this, *S. wallichianum* presents an excellent opportunity for further exploration.

In this study, we generated the chloroplast genome of *S. wallichianum* using bioinformatics tools. Specifically, we aim to assemble and annotate the plastome, resolve phylogenetic placement, and explore its conservation utility.

MATERIALS AND METHODS

Taxon sampling, DNA extraction, sequencing, assembly, and annotation

The *S. wallichianum* sample was collected from its native habitat at the Me Linh Station for Biodiversity in northern Vietnam, located at N 21°23.324', E 105°42.830' (Figure 1). The voucher specimen is deposited at the Hanoi Pedagogical University 2 with the voucher number Le243. To prevent DNA degradation, the fresh leaves of *S. wallichianum* were stored in silica gel at room temperature.

Total genomic DNA of *S. wallichianum* was extracted from silica gel-dried leaf samples using the CTAB method and purified with Qiagen Genomic-tip kit 13343 (Qiagen, Germany) to remove polysaccharides and protein contaminants (Doyle and Doyle 1987). Library construction and sequencing were performed by Benagen Technology Corporation Limited (Wuhan, China) following the DNBSEQ-T7 platform protocol. Raw sequencing reads were quality-filtered and adapter-trimmed using FASTP v0.23.2. After quality control, the clean reads were assembled into the chloroplast genome using GetOrganelle v1.7.5, which was selected for its high accuracy in reconstructing complete circular plastomes (Jin et al. 2020), and annotated using CPGAVAS2 (Shi et al. 2019). All annotation errors, as well as any ambiguous or unexpected features, were manually corrected using Apollo (Lewis et al. 2002). The circular structure of the chloroplast genome was drawn using OrganellarGenomeDRAW (OGDRAW) (Greiner et al. 2019).

Phylogenetic analysis

To construct the phylogenetic tree, we randomly selected plastomes of species closely related to Cervantesiaceae and *Scleropyrum* based on the studies of Nickrent et al. (2010) and Le (2018). In total, 13 plastomes of 13 species representing the families Opiliaceae, Thesiaceae, Cervantesiaceae, Nanodeaceae, Santalaceae, Viscaceae, and Amphorogynaceae were downloaded from the NCBI database (Table 1).

These species were selected to provide broad phylogenetic coverage within Santalales, enabling robust inference of the evolutionary position of *S. wallichianum*. The sequences were aligned using MAFFT v7.475 (Katoh and Standley 2013). *Schoepfia jasminodora* and *Schoepfia schreberi* were selected as the outgroup species. Gaps and poorly aligned regions were removed using trimAl (Capella-Gutiérrez et al. 2009). We then performed Maximum Likelihood (ML) analysis with RAxML v.8.0 (Stamatakis 2014), using default parameters and 1000 bootstrap replicates on the CIPRES Science Gateway Portal (<https://www.phylo.org>).

In addition, to compare the plastome information among species in Santalales, including *S. wallichianum*, we referred to data from 30 plastomes representing 30 species that encompass all life forms within Santalales according to Guo et al. (2020) (Table 2).

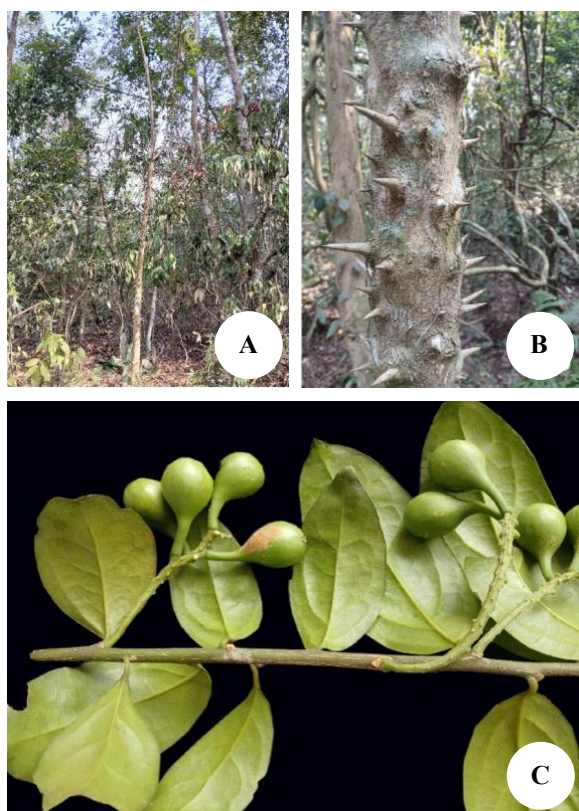


Figure 1. Morphology of *Scleropyrum wallichianum*: A. Habitat, B. Stem with thorns, and C. Fruit and leaves

Table 1. Voucher information and accession number of plastome used in this study

Species	Accession number
<i>Pyrularia edulis</i> A. DC.	MK675808
<i>Scleropyrum pentandrum</i> (Dennst.) Mabb.	ON881605
<i>Thesium chinense</i> Turcz.	MW846285
<i>Nanodea muscosa</i> Banks ex C.F.Gaertn.	OR712478
<i>Osyris wightiana</i> Wall. ex Wight	MK675807
<i>Osyris alba</i> L.	KT070882
<i>Santalum album</i> L.	MK675809
<i>Viscum ovalifolium</i> Wall. ex DC.	MK675811
<i>Viscum liquidambaricola</i> Hayata	MK675810
<i>Scleropyrum wallichianum</i> (Wight & Arn.) Arn.	PX255339
<i>Dendrotrophe varians</i> (Blume) Miq.	MF592987
<i>Cansjera rheedei</i> J.F.Gmel.	MN688989
<i>Schoepfia schreberi</i> J.F.Gmel.	OR712481
<i>Schoepfia jasminodora</i> Siebold & Zucc.	KX775962

RESULTS AND DISCUSSION

Chloroplast genome architecture of *S. wallichianum*

The study successfully sequenced the chloroplast genome of *S. wallichianum*. The plastome has been submitted to NCBI under the accession number PX255339 (Table 1). The assembly quality of the chloroplast genome was confirmed by sequencing depth analysis, with an average

coverage of 22,500× (minimum: 25×, maximum: 30,000×). In *S. wallichianum*, the complete chloroplast genome was assembled into a 149,066 bp circular DNA molecule. The complete chloroplast genome has a GC content of 38.30%. It consists of a Large Single-Copy (LSC) region of 85,000 bp, a Small Single-Copy (SSC) region of 18,500 bp, and two identical Inverted Repeat (IRa and IRb) regions, each measuring 22,783 bp.

Similar to other parasitic species within Santalales, the newly sequenced plastome of *S. wallichianum* exhibits the typical quadripartite structure, consisting of a pair of IRs, LSC, and SSC. As observed in other members of the genus *Scleropyrum*, the GC content of the IR regions was higher than that of the LSC and SSC regions, a pattern consistent with the molecular characteristics reported in the closely related species *P. edulis*. The plastome harbors 141 annotated genes, including 130 unique functional loci (80 protein-coding, 37 tRNA, and 8 rRNA genes), of which 12 are duplicated in the IR regions. Gene distribution analysis revealed that the LSC region accommodated 60 genes (45 protein-coding and 15 tRNA genes), while the SSC region contained 12 genes (10 protein-coding and 2 tRNA genes). Both IR regions each carried 29 genes, featuring 11 protein-coding genes (including *rpl2*, *rpl23*, *rps7*, and *yef2*), 8 tRNA genes (such as *trnA-UGC* and *trnI-GAU*), and 4 rRNA genes (*rrn16*, *rrn23*, *rrn4.5*, and *rrn5*). Structural analysis identified 17 intron-containing genes, with *clpP* and *yef3* exhibiting two introns each, and other genes like *atpF*, *ndhA*, *rpl16*, *rpoC1*, and *trnK-UUU* containing single introns. A unique trans-splicing arrangement was observed for *rps12*, where its 5'-end exon was situated in the LSC region, while duplicated 3'-end exons were located in the IR regions.

The plastome size and GC content of *S. wallichianum* are larger than several hemiparasitic species within Santalales in Table 2, such as *Arceuthobium sichuanense*, *Schoepfia jasminodora*, *Schoepfia fragrans*, *Taxillus chinensis*, *Taxillus nigrans*, *Taxillus sutchuenensis*, *Scurrula parasitica*, *Scurrula nothoxoides*, *Dendrophthoe pentandra*, *Loranthus tanakae*, *Tolypanthus maclurei*, *Helixanthera parasitica*, *Viscum crassulae*, *Viscum liquidambaricola*, *Viscum ovalifolium* (Guo et al. 2020). Although the plastome of *S. wallichianum* is larger than that of *P. edulis*, the two species share identical GC content (38.30%) (Table 2). In contrast, the plastome of *S. wallichianum* is smaller than that of *Ximenia americana*, a root-parasitic species (156,834 bp), and *Erythralum scandens*, an autotrophic species (156,154 bp) (Table 2).

Phylogenetic relationships

Phylogenetic analysis based on 14 complete plastomes in the ML tree shows the molecular relationships among species within Santalaceae sensu lato, with strong support for most clades (Figure 3). The five families—Santalaceae, Cervantesiaceae, Nonadeaceae, Amphorogynaceae, and Viscaceae formed a well-supported clade, and they are closely related (Figure 3). The species *S. wallichianum* was placed within the Cervantesiaceae family clade with strongly supported (Figure 3).

Table 2. Comparison of size and GC content (GC%) of complete plastomes, LSC, IR, and SSC regions between *Scleropyrum wallichianum* and other Santalales species. The plastome information of Santalales species refers to Guo et al. (2020)

Species	Lifeform	Plastome		LSC size (bp)	SSC size (bp)	IR size (bp)
		Size (bp)	GC (%)			
<i>Arceuthobium sichuanense</i> (H.S.Kiu) Hawksw. & Wiens	Aerial parasite	107,526	34.9	65,171	265	21,045
<i>Schoepfia jasminodora</i> Siebold & Zucc.	Aerial parasite	118,743	38.1	84,168	9,763	12,406
<i>Schoepfia fragrans</i> Wall.	Aerial parasite	120,188	38.1	85,643	9,783	12,381
<i>Taxillus chinensis</i> (DC.) Danser	Aerial parasite	121,363	37.3	70,357	6,082	22,462
<i>Taxillus nigrans</i> (Hance) Danser	Aerial parasite	121,419	37.4	70,181	6,100	22,569
<i>Scurrula parasitica</i> L.	Aerial parasite	121,750	37.2	70,270	6,106	22,687
<i>Dendrophthoe pentandra</i> (L.) Miq.	Aerial parasite	122,451	36.3	72,451	5,764	22,118
<i>Taxillus sutchuenensis</i> (Lecomte) Danser	Aerial parasite	122,562	37.3	70,630	6,102	22,915
<i>Loranthus tanakae</i> Franch. & Sav.	Aerial parasite	123,397	36.9	69,522	7,723	23,076
<i>Tolypanthus maclurei</i> (Merr.) Danser	Aerial parasite	123,581	36.8	72,952	6,259	22,185
<i>Scurrula notothixoides</i> (Hance) Danser	Aerial parasite	123,810	37.3	71,448	6,160	23,101
<i>Helixanthera parasitica</i> Lour.	Aerial parasite	124,881	36.5	73,043	6,334	22,752
<i>Malania oleifera</i> Chun & S.K.Lee	Aerial parasite	125,050	38.2	76,387	15	24,324
<i>Viscum crassulae</i> Eckl. & Zeyh.	Aerial parasite	126,064	36.4	73,226	8,628	22,105
<i>Macrosolen tricolor</i> (Lecomte) Danser	Aerial parasite	126,617	37.6	71,893	5,320	24,702
<i>Macrosolen</i> sp. (Blume) Rchb.	Aerial parasite	128,459	37.3	73,347	9,192	22,960
<i>Viscum liquidambaricola</i> Hayata	Aerial parasite	128,601	36.1	73,831	8,688	23,041
<i>Viscum coloratum</i> (Kom.) Nakai	Aerial parasite	128,746	36.3	73,686	8,630	23,215
<i>Viscum album</i> L.	Aerial parasite	128,921	36.4	73,893	8,632	23,198
<i>Viscum ovalifolium</i> DC.	Aerial parasite	129,465	36.1	74,348	8,711	23,203
<i>Viscum yunnanense</i> H.S.Kiu	Aerial parasite	130,721	35.8	75,844	9,089	22,894
<i>Viscum minimum</i> Harv.	Aerial parasite	131,016	36.2	75,814	9,014	23,094
<i>Pyricularia edulis</i> A.DC.	Root parasite	132,808	38.3	74,811	8,901	24,548
<i>Dendrotrophe varians</i> (Blume) Miq.	Aerial parasite	140,666	37.8	81,684	10,870	24,056
<i>Santalum album</i> L.	Root parasite	144,101	38.0	83,802	11,277	24,511
<i>Osyris alba</i> L.	Root parasite	147,253	37.7	84,601	13,972	24,340
<i>Champereia manillana</i> (Blume) Merr.	Root parasite	147,461	37.4	83,505	7,806	28,075
<i>Osyris wightiana</i> Wall. ex Wight	Root parasite	147,544	37.6	84,569	14,081	24,447
<i>Scleropyrum wallichianum</i> (Wight & Arn.) Arn.	Root parasite	149,066	38.3	85,000	18,500	45,566
<i>Erythralum scandens</i> Blume	Autotroph	156,154	38.0	84,799	18,567	26,394
<i>Ximения americana</i> L.	Root parasite	156,834	36.8	87,816	3,636	32,691

Discussion

Comparative plastome characteristics among Santalales lineages and inferences of parasitic lifestyles

Chloroplast genomes provide comprehensive information, due to their non-recombinant and uniparentally inherited nature, making them useful tool for ecological, evolutionary, and conservation studies (Baasanmunkh et al. 2025). *S. wallichianum* exists in a small population and is likely to decline under the impacts of climate change and human activity. The genetic variation of *S. wallichianum* can be assessed based on chloroplast genome data, which is crucial for the conservation of endangered wild plants.

Guo et al. (2020) suggested that the transition from autotrophy to parasitism is often accompanied by a reduction in plastome size, a pattern observed in parasitic species of the sandalwood order (Santalales) (Table 1). This plastome reduction is primarily due to the loss of non-coding regions and reflects an adaptation to the parasitic lifestyle, where plants rely on their hosts for nutrients. *S. wallichianum* is a root-parasitic species, and its relatively large plastome suggests a life form capable of partially sustaining itself by acquiring nutrients from the environment while also obtaining supplementary nutrition from its host (Wu et al. 2024).

Erythralum scandens (autotrophic species), *S. wallichianum* (root-parasitic species), and *X. americana*

(root-parasitic species) have the largest plastomes in size, while the plastomes of the remaining species show a reduction in size. The reduction in plastome size is mainly observed in the LSC and SSC regions (Table 2). Hu et al. (2015), Li et al. (2020), and Maurya et al. (2025) suggested that the LSC and SSC regions contain many genes involved in photosynthesis and translation, and these regions evolve at a faster rate than the Inverted Repeat (IR) region. Thus, the reduced plastome size of aerial parasitic species indicates their adaptation to a parasitic lifestyle and greater dependence on their hosts.

In contrast, *S. wallichianum* and *X. americana* have plastomes similar in size to those of *E. scandens*, which may reflect retention of additional genes or noncoding regions, potentially associated with its hemiparasitic lifestyle and structural stability within the chloroplast genome. This result suggested that they maintain good autotrophic ability with little dependence on their hosts. Additionally, evidence of plastome loss in holoparasites, such as *Cytinus hypocistis* (Roquet et al. 2016), *Balanophora laxiflora* (Su et al. 2019), *Ombrophytum subterraneum* and *Lophophytum leandri* (Ceriotti et al. 2021), and *Sarcophyte sanguinea* and *Thonningia sanguinea* (Kim et al. 2023), demonstrates the high degree of adaptation to non-photosynthetic parasitism in this plant group.

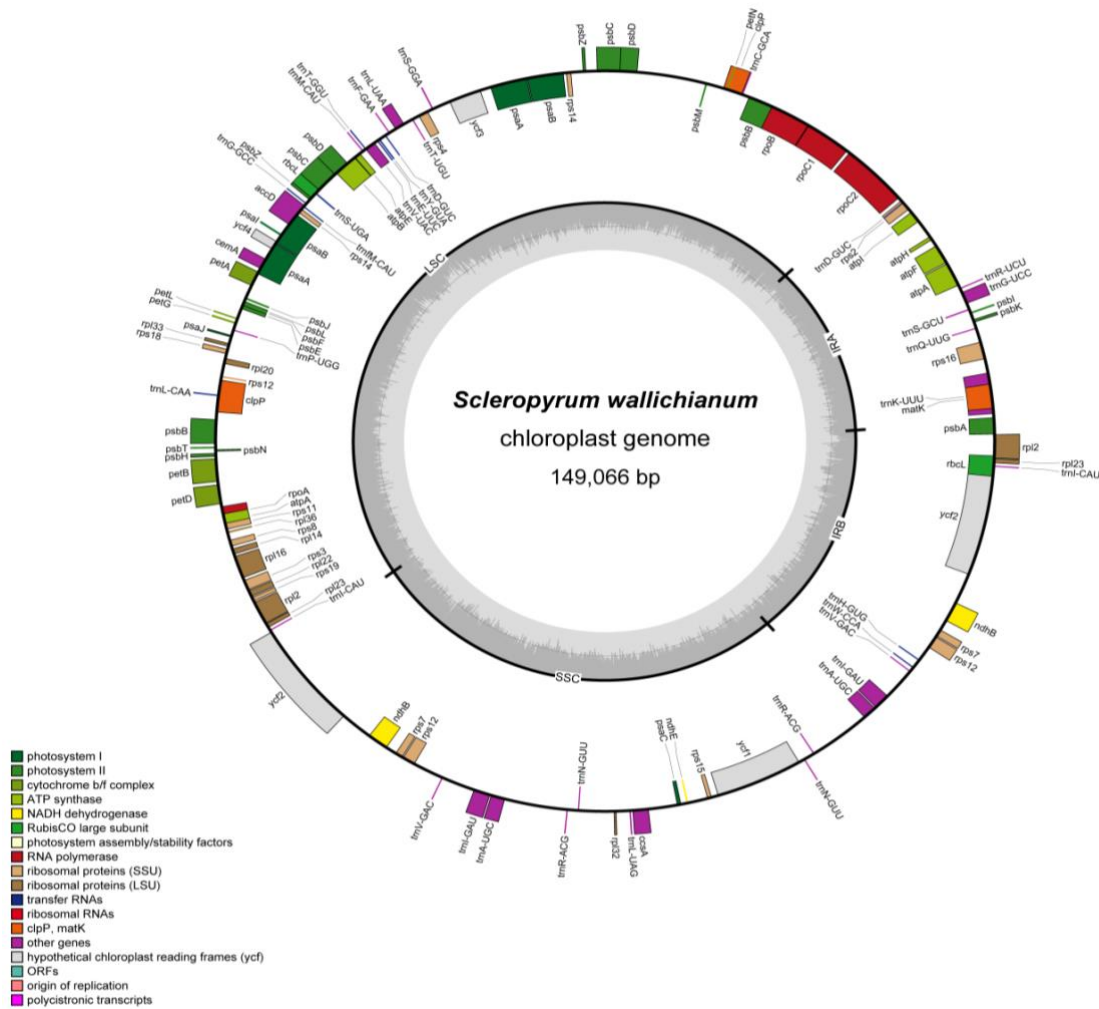


Figure 2. Chloroplast genome map of *Scleropyrum wallichianum*. The map illustrates the typical quadripartite structure of the chloroplast genome, comprising a Large Single-Copy (LSC) region, a Small Single-Copy (SSC) region, and two Inverted Repeats (IRs). Genes located outside the main circle are transcribed clockwise, while those inside are transcribed counterclockwise. Genes are color-coded based on functional categories, such as photosynthesis, ribosomal proteins, and RNA polymerase. The inner gray circle represents the GC content variation across the genome. The complete chloroplast genome is 149,066 bp in length. This figure was generated using OGDRAW v1.3.1 (OrganellarGenomeDRAW; Lohse et al. 2013); <https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>

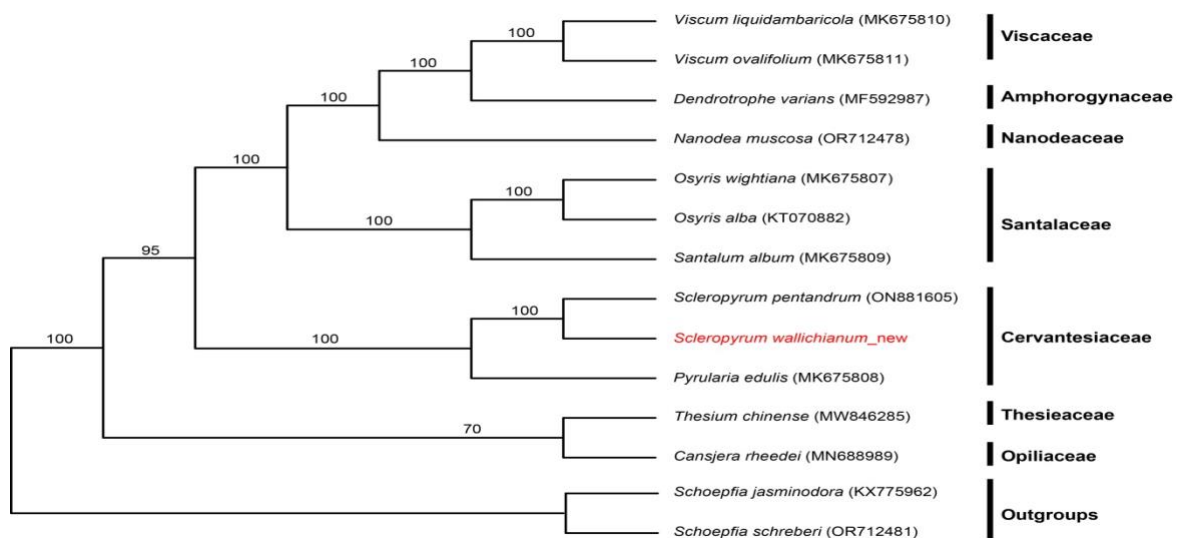


Figure 3. Maximum Likelihood (ML) phylogenetic tree of *Scleropyrum wallichianum* and related species based on complete chloroplast genome sequences. The tree was constructed using RAxML v8.0. Bootstrap support values from the ML analysis are shown above the branches

Phylogenetic relationships

The molecular phylogenetic relationships inferred in this study are congruent with previous studies on the genus *Scleropyrum* (Der and Nickrent 2008; Le et al. 2019, 2024; Zhou et al. 2019). Our results support the monophyly of Cervantesiaceae and other families within Santalaceae sensu lato. *S. wallichianum*, *S. pentandrum*, and *P. edulis* are closely related within the family Cervantesiaceae (Figure 3). Li et al. (2020) and Maurya et al. (2025) noted that the LSC and SSC regions of the plastome contain a greater number of genes, making them valuable for phylogenetic analysis, species identification, and evolutionary studies. The findings from the plastome of *S. wallichianum* and other species in Santalales highlight the importance of generating genomic data for members of the Sandalwood order to facilitate future phylogenetic, evolutionary, and conservation research. The complete chloroplast genome of *S. wallichianum* provides a foundational genomic resource for conservation genetics, enabling assessment of population genetic diversity and structure. Such information can inform management strategies for endangered populations and guide targeted conservation measures. Integrating plastome data with ecological and reproductive studies will further support conservation planning and sustainable management of this economically and ecologically important species.

From a conservation perspective, the plastome sequence enhances our ability to identify and differentiate populations, assess genetic diversity, and monitor lineage divergence across the species' distribution range. Such genomic tools are essential for designing effective conservation strategies, particularly for taxa that experience habitat fragmentation and overexploitation due to economic uses, such as seed oil extraction. Moreover, chloroplast markers can be developed from the genome to support DNA barcoding and species authentication in trade and restoration programs (Li et al. 2019). Comparative plastome analyses among Cervantesiaceae and related families will also help reveal the evolutionary adaptations associated with parasitism and environmental tolerance, contributing to the understanding of resilience mechanisms under changing climates (Wicke et al. 2011). The sequencing of the complete chloroplast genome of *S. wallichianum*, a member of the family Cervantesiaceae, represents a significant advancement for both evolutionary research and conservation biology.

In conclusion, this study presents the first complete chloroplast genome of *S. wallichianum* (Cervantesiaceae), providing a valuable genomic resource for phylogenetic inference and conservation genetics. The plastome (149,066 bp; GC 38.3%) displays a typical quadripartite structure and strong gene conservation within Santalales, confirming *S. wallichianum* as a distinct clade closely related to Opiliaceae. These findings contribute to a better understanding of the evolutionary patterns and adaptive features of hemiparasitic lineages. Nevertheless, the study is limited by its single-sample design and lack of nuclear and mitochondrial genome comparisons, which constrain insights into intraspecific variation and overall genome evolution. Future research should include population-level chloroplast variation, multi-genomic integration, and

comparative analyses among Cervantesiaceae taxa. Such approaches will deepen our understanding of evolutionary adaptation, taxonomic relationships, and conservation priorities for parasitic groups within Santalales.

ACKNOWLEDGEMENTS

This research was funded by the Vietnam Ministry of Education and Training under grant number B.2024-SP2-05. We thank Prof. Miao Sun from Huazhong Agricultural University, China, for his help in the initiation stage of this study.

REFERENCES

- Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. *Nat Rev Genet* 11: 697-709. DOI: DOI: 10.1038/nrg2844.
- Baasanmunkh S, Nyamgerel N, Oyuntsetseg B, Munkhtulga D, Erst A, Kubentayev S, Bayarmaa G-A, Yoon JW, Choi HJ. 2025. The complete chloroplast genome of *Tulipa uniflora* (Liliaceae) in Mongolia. *Mitochondrial DNA B Resour* 10 (3): 257-261. DOI: 10.1080/23802359.2025.2468748.
- Cardinale BJ, Duffy JE, Gonzalez A, et al. 2012. Biodiversity loss and its impact on humanity. *Nature* 486: 59-67. DOI: DOI: 10.1038/nature11148.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972-1973. DOI: 10.1093/bioinformatics/btp348.
- Ceriotti LF, Roulet ME, Sánchez-Puerta MV. 2021. Plastomes in the holoparasitic family Balanophoraceae: Extremely high AT content, severe gene content reduction, and two independent genetic code changes. *Mol Phylogenet Evol* 162: 107208. DOI: 10.1016/j.ympev.2021.107208.
- Choi Y-R, Kim S-C, Kim T-H, Ha Y-H, 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 B Resour* 9 (8): 1005-1009. DOI: 10.1080/23802359.2024.2387262.
- Coissac E, Hollingsworth PM, Lavergne S, Taberlet P. 2016. From barcodes to genomes: Extending the concept of DNA barcoding. *Mol Ecol* 25: 1423-1428. DOI: 10.1111/mec.13549.
- Daniell H, Lin CS, Yu M, Chang WJ. 2016. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol* 17: 134. DOI: 10.1186/s13059-016-1004-2.
- Der JP, Nickrent DL. 2008. A molecular phylogeny of Santalaceae (Santalales). *Syst Bot* 33: 107-116. DOI: 10.1600/036364408783887438.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 9: 11-15.
- 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.
- Guo X, Liu C, Zhang G, Su W, Landis JB, Zhang X, Wang H, Ji Y. 2020. The complete plastomes of five hemiparasitic plants (*Osyris wightiana*, *Pyrrularia edulis*, *Santalum album*, *Viscum liquidambaricolum*, and *V. ovalifolium*): Comparative and evolutionary analyses within Santalales. *Front Genet* 11: 597. DOI: 10.3389/fgene.2020.00597.
- Hu S, Sablok G, Wang B, Qu D, Barbaro E, Viola R, Li M, Varotto C. 2015. Plastome organization and evolution of chloroplast genes in *Cardamine* species adapted to contrasting habitats. *BMC Genomics* 16 (1): 306. DOI: 10.1186/s12864-015-1498-0.
- Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 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.
- Jin L, Liu J-J, Li Q-M, Lin L-X, Shao X-N, Xiao T-W, Li B-H, Mi X-C, Ren H-B, Zhu Y, Qiao X-J, Lian J-Y, Wang X-G, Du H, Jin G-Z, Svenning J-C, Hao G, Ge X-J. 2023. Stronger latitudinal phylogenetic

- patterns in woody angiosperm assemblages with higher dispersal abilities in China. *J Biogeogr* 51: 269-279. DOI: 10.1111/jbi.14746.
- 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.
- Kim W, Lautenschläger T, Bolin JF, Rees M, Nzuzi A, Zhou R, Wanke S, Jost M. 2023. Extreme plastomes in holoparasitic Balanophoraceae are not the norm. *BMC Genomics* 24 (1): 330. DOI: 10.1186/s12864-023-09422-1.
- Le CT, Do TB, Nguyen TL, Nguyen TMA, Dang TA. 2019. Exploring the phylogenetic relationship of Santalaceae genera based on intensive molecular data. *Acad J Biol* 41 (2se1-2se2): 221-228.
- Le CT, Nguyen VD, Ha MT, Pham TBH, Hoang TPN, Pham TH, Liu B. 2024. *Pyrularia*: A newly recorded genus of Cervantesiaceae from Vietnam. *Phytotaxa* 676: 194-198. DOI: 10.11646/phytotaxa.676.2.7.
- Le CT. 2018. Phylogeny, biogeography and diversification of Santalales. [Dissertation]. University of Chinese Academy of Sciences, Beijing, China.
- Lewis SE, Searle SMJ, Harris N, Gibson M, Lyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby MA, Kaminker JS, Matthews BB, Prochnik SE, Smithy CD, Tupy JL, Rubin GM, Misra S, Mungall CJ, Clamp ME. 2002. Apollo: A sequence annotation editor. *Genome Biol* 3 (12): research0082. DOI: 10.1186/gb-2002-3-12-research0082.
- Li L, Jiang Y, Liu Y, Niu Z, Xue Q, Liu W, Ding X. 2020. The Large Single-Copy (LSC) region functions as a highly effective and efficient molecular marker for accurate authentication of medicinal *Dendrobium* species. *Acta Pharm Sin B* 10 (10): 1989-2001. DOI: 10.1016/j.apsb.2020.01.012.
- Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. 2019. Plant DNA barcoding: from gene to genome. *Biol Rev* 90: 157-166. DOI: 10.1111/brv.12104.
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression. *Nucleic Acids Res* W1: W575-W581. DOI: 10.1093/nar/gkt289.
- Maurya S, Sukhramani G, Lee C, Jo S, Kim J, Jeong EJ, Zamora NA, Garcia R, Zhang Z, Choi S, Choudhary RK, Kim S-Y. 2025. Comparative plastome evaluation, phylogenomic analysis, and DNA signatures of medicinally important malvaceous genera and their adulterants. *Sci Rep* 15 (1): 24285. DOI: 10.1038/s41598-025-07744-6.
- Nickrent DL. 2011. Santalales (Including Mistletoes). *Encyclopedia of Life Sciences* 1-6. DOI: 10.1002/9780470015902.a0003714.pub2.
- Nickrent DL, Malécot V, Vidal-Russell R, Der JP. 2010. A revised classification of Santalales. *Taxon* 59 (2): 538-558. DOI: 10.1002/tax.592019.
- Peruzzi L. 2016. A new infrafamilial taxonomic setting for Liliaceae, with a key to genera and tribes. *Plant Biosyst* 150 (6): 1341-1347. DOI: 10.1080/11263504.2015.1115435.
- Pimm SL, Jenkins CN, Abell R, et al. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344: 1246752. DOI: 10.1126/science.1246752.
- Roquet C, Coissac É, Cruaud C, Boleda M, Boyer F, Alberti A, Gielly L, Taberlet P, Thuiller W, Van Es J, Lavergne S. 2016. Understanding the evolution of holoparasitic plants: The complete plastid genome of the holoparasite *Cytinus hypocistis* (Cytinaceae). *Ann Bot* 118 (5): 885-896. DOI: 10.1093/aob/mcw135.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res* 47 (W1): W65-W73. DOI: 10.1093/nar/gkz345.
- Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30 (9): 1312-1313. DOI: 10.1093/bioinformatics/btu033.
- Su H-J, Barkman TJ, Hao W, Jones SS, Naumann J, Skippington E, Wafula EK, Hu J-M, Palmer JD, dePamphilis CW. 2019. Novel genetic code and record-setting AT-richness in the highly reduced plastid genome of the holoparasitic plant *Balanophora*. *Proc Natl Acad Sci USA* 116 (3): 934-943. DOI: 10.1073/pnas.1816822116.
- Wicke S, Schneeweiss GM, Müller KF, dePamphilis CW, Quandt D. 2011. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol Biol* 76: 273-297. DOI: 10.1007/s11103-011-9762-4.
- Wu J, Zhang L, Ma X, Fu X, Chen F, Jiao Y, Liu J, Wu S. 2024. The evolutionary significance of whole genome duplications in oil biosynthesis of oil crops. *Hortic Res* 11 (7): uhae156. DOI: 10.1093/hr/uhae156.
- Xia N, Gilbert M G. 2003. Santalaceae. In: Wu ZY, Raven PH, Hong DY (eds). *Flora of China*. Science Press, Beijing and Missouri Botanical Garden Press, St. Louis.
- Zhou Z, Hu J-J, Wen J, Sun H. 2019. Morphometric, phylogenetic, and biogeographic analyses of *Pyrularia* (Santalales), a parasitic disjunct lineage between eastern Asia and eastern North America. *Taxon* 68 (1): 47-71. DOI: 10.1002/tax.12021.