



OPEN ACCESS

EDITED BY

Wanli Zhao,
Jiangsu Province and Chinese Academy
of Sciences, China

REVIEWED BY

Roohaida Othman,
Universiti Kebangsaan Malaysia, Malaysia
Wenxuan Dong,
Shenyang Agricultural University, China
Hoang Dang Khoa Do,
Nguyen Tat Thanh University, Vietnam

*CORRESPONDENCE

Diana Prameswari
✉ dian074@brin.go.id

RECEIVED 21 January 2026

REVISED 22 April 2026

ACCEPTED 28 April 2026

PUBLISHED 19 May 2026

CITATION

Sudarmono S, Nugroho A,
Damayanti RU, Sudomo A, Lailaty IQ,
Diantina S, Nugraheni YMM, Palupi KD
and Prameswari D (2026) A genomic
resource for medicinal plant research:
chloroplast genome assembly and
annotation of *Helicteres isora* L..
Front. Plant Sci. 17:1792650.
doi: 10.3389/fpls.2026.1792650

COPYRIGHT

© 2026 Sudarmono, Nugroho,
Damayanti, Sudomo, Lailaty, Diantina,
Nugraheni, Palupi and Prameswari. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

A genomic resource for medicinal plant research: chloroplast genome assembly and annotation of *Helicteres isora* L.

Sudarmono Sudarmono¹, Aditya Nugroho²,
Ratna Uli Damayanti², Aris Sudomo³, Intani Quarta Lailaty²,
Surya Diantina², Yosephin Martha Maria Anita Nugraheni²,
Kartika Dyah Palupi⁴ and Diana Prameswari^{2*}

¹Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Bogor, West Java, Indonesia, ²Research Center for Applied Botany, National Research and Innovation Agency (BRIN), Bogor, West Java, Indonesia, ³Research Center for Ecology, National Research and Innovation Agency (BRIN), Bogor, West Java, Indonesia, ⁴Research Center for Pharmaceutical Ingredients and Traditional Medicine, National Research and Innovation Agency (BRIN), Bogor, West Java, Indonesia

KEYWORDS

chloroplast, genome, *Helicteres isora*, Helicteroideae, Indonesia

Introduction

Rapid advances in plant genomics have opened up tremendous opportunities for understanding the genetic basis, evolution, and adaptation of various plant species, including medicinal plants. Genomic data is now an important foundation for studies in systematics, phylogenetics, conservation, and the sustainable use of biological resources. Medicinal plants have the potential to be developed as new model systems (emerging models) because they have a unique genomic architecture and regulatory pathways that are important for understanding secondary metabolism and adaptive evolution (Ray et al, 2022). The shift from laboratory model systems to non-model species in natural environments reflects the need to understand natural variation and molecular mechanisms in a more realistic ecological context. (Richards et al, 2017).

The chloroplast genome is an important component in the study of diversity, evolution, and the application of plant genetic techniques. (Daniell et al, 2016). In general, the chloroplast genome has a quadripartite structure consisting of large single copy (LSC), small single copy (SSC), and two inverted repeats (IR) regions, with a relatively conserved gene composition. Recent studies show that chloroplast DNA (cpDNA) is not only circular, but can also exist as a branched linear structure, reflecting a more dynamic and complex chloroplast genetic system than previously understood (Dobrogojski et al, 2020). A deep understanding of the chloroplast genome has fundamental and applied value, especially in plant biotechnology, although chloroplast genome modification still faces technical and biological challenges. (Olejniczak et al, 2016).

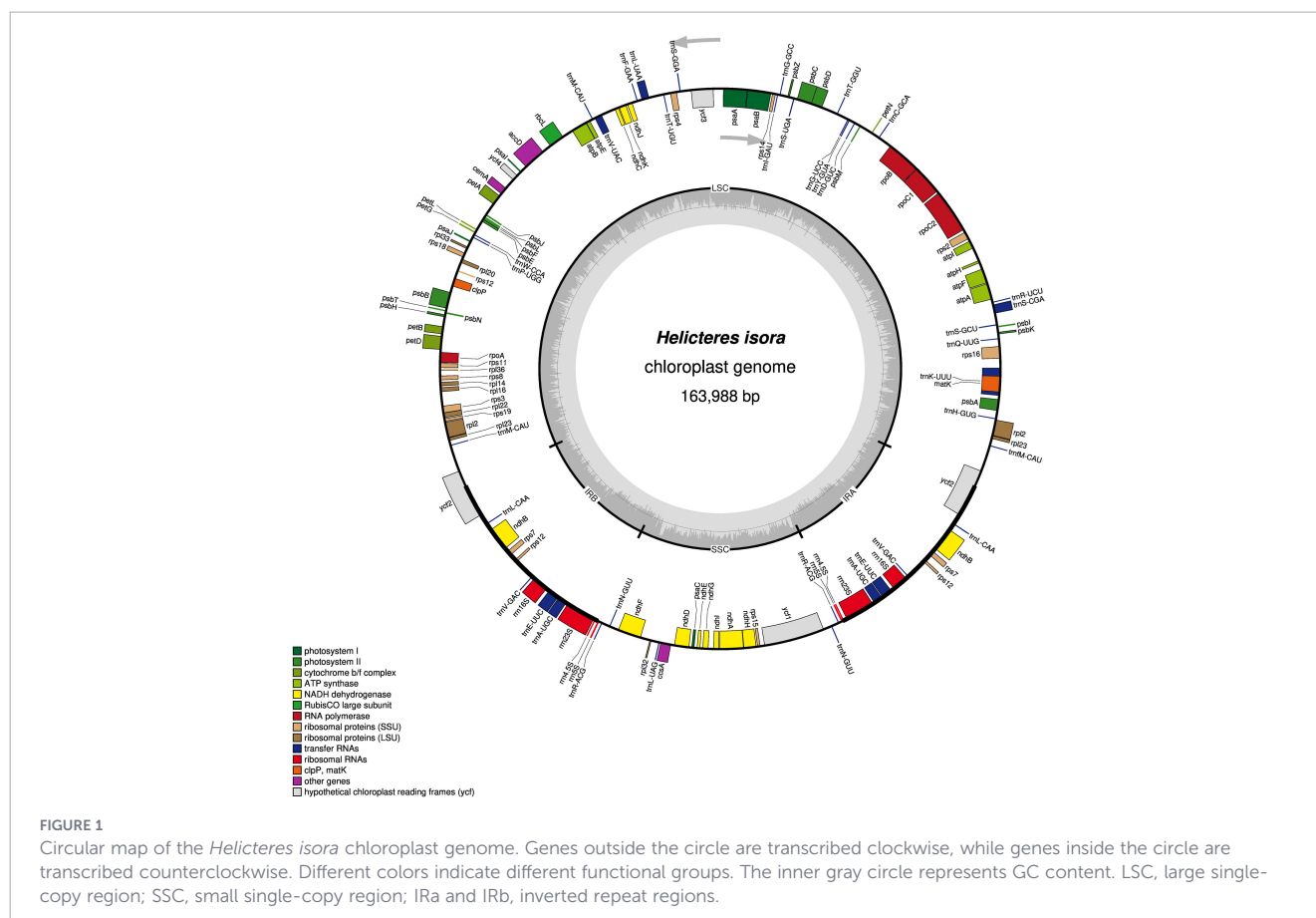
In the context of plant identification and authentication, chloroplast genomes play an important role in the development of DNA barcoding, especially for medicinal plants.

However, the absence of a universal barcode locus for all plant groups makes the selection of informative chloroplast markers a crucial aspect in phytomedicine authentication and biodiversity conservation (Alam et al, 2024). Information obtained from chloroplast genomes, such as gene organization, codon usage patterns, and sequence variation, has been widely used for high-resolution phylogenetic analysis, molecular marker development, and comparative studies between species. In the context of medicinal plants, the availability of complete chloroplast genomes is becoming increasingly important as it can support species authentication, tracking evolutionary relationships, and exploring genetic diversity relevant to conservation and pharmacological utilization.

Screw tree (*Helicteres isora*), known locally as kayu ules, is one of the medicinal plant species that grows naturally and is still rarely cultivated. According to the 2024 IUCN Red List, this species is categorized as Least Concern. Morphologically, *H. isora* belongs to the Malvaceae family and is an upright shrub with a height of about 1–2.5 m, with red flowers (Kanthale and Biradar, 2017). The fruit is arranged in a spiral or twisted pattern, which is the basis for its common name, screw tree. (Kumbhani et al, 2017). Various parts of this plant, including the bark, roots, leaves, seeds, and fruit, have been reported to have medicinal properties, including for treating diarrhea (bark), diabetes, weakness in newborns, and ear infections (fruit), dysentery (seeds), and scabies (leaves). These pharmacological activities are related to the content of important bioactive

compounds, such as betulinic acid, β -tocopherol, coumarin, β -sitosterol, saponins, flavonoids, tannins, and alkaloids (Dayal et al, 2015; Kanthale and Biradar, 2017; Meena et al, 2018).

Despite its high pharmacological value, *Helicteres isora* does not yet have a publicly available annotated chloroplast genome resource. Currently, only one chloroplast genome sequence from a closely related species, *Helicteres hirsuta* (Nguyen et al, 2024), has already been reported and provides an important comparative reference for genome annotation and structural characterization in this genus. The availability of plastome data from related taxa highlights the value of comparative chloroplast genomics for improving genome annotation accuracy and for supporting evolutionary and phylogenetic studies within *Helicteres*. However, the lack of comprehensive genomic information has limited evolutionary and ecological studies, particularly in understanding natural variation and local adaptation on a broader genomic scale (Song & Mitchell-Olds, 2011). Therefore, this study aimed to assemble, annotate, and characterize the complete chloroplast genome of *H. isora* L. as a basic genomic resource for medicinal plant research. This study focused on the analysis of plastome architecture, gene composition, and genome organization. The resulting chloroplast genome data are expected to serve as an important reference for future studies in systematics, evolutionary biology, genetic conservation, DNA barcoding, and bioprospecting, as well as a basis for formulating more effective conservation and management strategies for this valuable medicinal plant.



Methods

Plant materials

Fresh leaves were collected from a single individual of *H. isora* collected in Bosen village, North Mollo Sub District, South Timor Tengah District, East Nusa Tenggara Province, Indonesia (10° 09'34.95" S, 123°35'03.32" E).

DNA extraction

Fresh leaf tissue of *Helicteres isora* was used for genomic DNA extraction. Total genomic DNA was isolated using the SEPA Plant DNA Isolation Reagent Kit (KIT-9101) following the manufacturer's instructions. The concentration and purity of the extracted DNA were initially assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific) by measuring the A260/280 and A260/230 ratios. DNA integrity was evaluated by 1% agarose gel electrophoresis. Accurate DNA quantification was performed using the Qubit dsDNA High Sensitivity Assay Kit (Thermo Scientific). Further assessment of DNA integrity was carried out using the Agilent 4150 TapeStation.

Library preparation and sequencing

High-quality genomic DNA was used as input for library preparation using the xGen DNA Library Prep EZ UNI Kit (Integrated DNA Technologies, IDT). Genomic DNA was enzymatically fragmented to obtain the appropriate insert size, followed by ligation with Illumina-compatible adapters. PCR amplification was subsequently performed to generate full-length adapters containing unique index sequences for each library. The quality and quantity of the prepared libraries were evaluated using a Qubit Fluorometer, Agilent TapeStation, and quantitative PCR (qPCR). Paired-end sequencing (2 × 150 bp) was conducted on the Illumina NextSeq 2000 platform by PT. Genetika Science Indonesia. Base calling, demultiplexing, and FASTQ file generation were performed using Illumina DRAGEN BCL Convert v4.2.7.

Chloroplast genome assembly and annotation

Raw reads were evaluated for quality using FASTQC v0.11.8 (Andrews et al, 2010). Adapter sequences, low-quality bases (quality score <30), nucleotide bias at the 5' and 3' ends, and potential contaminant sequences were removed using Trimmomatic v0.39. The trimming process employed the following parameters: TruSeq3-PE.fa:2:30:10, SLIDINGWINDOW:4:28, LEADING:28, TRAILING:28, and MINLEN:20 (Bolger et al, 2014). The high-quality filtered reads were subsequently assembled using GetOrganelle v1.7.7.1 (Jin et al, 2020). The annotation was carried out using CPGAVAS2 (Shi et al, 2019), with the chloroplast genome of *Helicteres hirsuta* (accession number MW029459) used as the reference sequence. Challenges encountered during annotation, including ambiguous gene regions and uncertain gene boundaries, were resolved through manual curation using Unipro UGENE

v45.1 (Okonechnikov et al, 2012) and NCBI Genomic Workbench v3.8.2 (Kuznetsov and Bollin, 2021). To ensure sequence completeness, the assembled chloroplast genome was verified to contain no ambiguous nucleotides (N) and to encode all 21 amino acids using Unipro UGENE v45.1. Genes lacking a start codon were manually corrected using the editing functions in NCBI Genomic Workbench v3.8.2. Finally, visualization of the circular chloroplast genome was generated using Organellar Genome DRAW (OGDRAW) via the MPI-MP Chlorobox platform (Greiner et al, 2019).

Description of the chloroplast genome dataset

The dataset comprises the complete chloroplast genome sequence of *Helicteres isora* assembled into a single circular molecule. The chloroplast genome has a total length of 163,988 bp and exhibits a quadripartite organization consisting of a large single-copy (LSC) region of 91,679 bp, a small single-copy (SSC) region of 19,789 bp, and two inverted repeat (IR) regions of 26,260 bp each (Figure 1). Base composition statistics indicate that adenine (31.10%) and thymine (32.20%) are more frequent than cytosine (18.60%) and guanine (18.00%), resulting in an overall GC content of 37.00%. Regional GC content values were calculated for the LSC (34.47%), SSC (37.78%), and IR (42.28%) regions to support comparative analyses of plastome composition. This GC enrichment in the IR regions is consistent with the presence of GC-rich rRNA genes and reflects a conserved compositional pattern commonly observed in chloroplast genomes of flowering plants (Lestari et al, 2024; Nugroho et al, 2025; Nuroniah et al, 2025).

A total of 113 unique genes were annotated in the chloroplast genome of *H. isora*, comprising 79 protein-coding genes (PCGs), 30 transfer RNA (tRNA) genes, and four ribosomal RNA (rRNA) genes (Table 1). These genes were classified into functional categories associated with photosynthesis, genetic self-replication, and other essential chloroplast functions. Genes involved in photosynthesis constitute a substantial proportion of the plastome and include those encoding subunits of photosystem I (*psa* genes), photosystem II (*psb* genes), the cytochrome b/f complex (*pet* genes), ATP synthase (*atp* genes), and NADH dehydrogenase (*ndh* genes), as well as the large subunit of Rubisco (*rbcL*). The presence and arrangement of these genes indicate a highly conserved photosynthetic apparatus in *H. isora*, consistent with other Malvaceae plastomes. Genes related to genetic self-replication include ribosomal protein genes of both the large (*rpl*) and small (*rps*) subunits, DNA-dependent RNA polymerase genes (*rpoA-rpoC2*), and a complete set of tRNA genes sufficient for chloroplast protein synthesis.

Gene duplication associated with the inverted repeat regions was documented in the dataset. Several genes are present in two copies due to their localization within the IR regions, including the rRNA genes *rrn16S*, *rrn23S*, *rrn5S*, and *rrn4.5S*; the tRNA genes *trnA-UGC*, *trnE-UUC*, *trnL-CAA*, *trnM-CAU*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*; and the protein-coding genes *ndhB*, *rpl2*,

TABLE 1 List of genes in the *Helicteres isora* chloroplast genome.

Functional category	Group of gene	Gene
Self-replication	rRNA	<i>rrn16S^d, rrn23S^d, rrn5S^d, rrn4.5S^d</i>
	tRNA	<i>trnA-UGC^{d*}, trnC-GCA, trnD-GUC, trnE-UUC^{d*}, trnF-GAA, trnG-GCC, trnG-UCC, trnH-GUG, trnI-GAU, trnK-UUU*, trnL-CAA^d, trnL-UAA*, trnL-UAG, trnM-CAU, trnM-CAU^d, trnN-GUU^d, trnP-UGG, trnQ-UUG, trnR-ACG^d, trnR-UCU, trnS-CGA, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC^d, trnV-UAC*, trnW-CCA, trnY-GUA</i>
	Subunits of ATP synthase	<i>atpA, atpB, atpE, atpF*, atpH, atpI</i>
	Large subunit of ribosome	<i>rpl14, rpl16, rpl2^{d*}, rpl20, rpl22, rpl23^d, rpl32, rpl33, rpl36</i>
	Small subunit of ribosome	<i>rps11, rps12^{d**}, rps14, rps15, rps16*, rps18, rps19, rps2, rps3, rps4, rps7^d, rps8</i>
	Subunits of NADH-dehydrogenase	<i>ndhA*, ndhB^{d*}, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
Photosynthesis	Subunits of cytochrome b/f complex	<i>petA, petB, petD*, petG, petL, petN</i>
	Subunits of photosystem I	<i>psaA, psaB, psaC, psaI, psaJ</i>
	Subunits of photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ, ycf3**</i>
	Subunit of rubisco	<i>rbcL</i>
	DNA-dependent RNA polymerase	<i>rpoA, rpoB, rpoC1*, rpoC2</i>
	Subunit of Acetyl-CoA-carboxylase	<i>accD**</i>
	c-type cytochrome synthesis gene	<i>ccsA</i>
Other function	Envelope membrane protein	<i>cemA</i>
	Protease	<i>clpP</i>
	Maturation	<i>matK</i>
Conserved open reading frames	Conserved open reading frames	<i>ycf1**, ycf2^{d*}, ycf4</i>

d, gene duplication; *, single intron; **, double intron.

rpl23, *rps7*, *rps12*, and *ycf2*. In addition, intron-containing genes were identified and annotated. Single introns were recorded in genes such as *atpF*, *ndhA*, *ndhB*, *petD*, *rpoC1*, and *rps16*, while genes including *ycf1*, *ycf3*, *rps12*, and *accD* contain two introns. The *rps12* gene exhibits a trans-splicing configuration, with exons distributed across different regions of the genome. Structural features of the *clpP* gene, including exon–intron organization, were also documented as part of the annotation output. All annotated genes, their functional categories, copy number, and intron status are summarized in Table 1. The complete chloroplast genome sequence and associated annotation files are publicly available through the NCBI GenBank database and are intended to support reuse in plastome annotation, molecular marker development, comparative genomics, and other data-driven studies involving medicinal plants.

Data availability statement

The dataset presented in this study is available in the NCBI GenBank database. The complete chloroplast genome sequence of *Helicteres isora* has been deposited under accession number PX570609.

Author contributions

SS: Formal analysis, Methodology, Writing – original draft. AN: Writing – review & editing, Data curation. RD: Writing – review & editing, Formal analysis. AS: Writing – review & editing, Supervision, Resources. IL: Software, Writing – review & editing. SD: Investigation, Supervision, Writing – review & editing. YN: Writing – review & editing, Visualization. KP: Validation, Writing – review & editing. DP: Supervision, Writing – review & editing, Project administration, Conceptualization, Funding acquisition.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was also supported by Riset Inovasi Indonesia Maju (RIIM) 7th, with funding provided by LPDP (Ministry of Finance, Republic of Indonesia), under Contract No: B_3864/II.7.5/KS.00/4/2025. Additional support was provided by the LPDP RIIM Terrestrial Biodiversity Expedition under Contract Nos. 7501/II.7.5/FR.06.00/11/2024 and B-7766/III.5/FR.06.00/11/2024.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial

intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Alam, A., Usman, M., Parveen, B., and Ahmad, S. (2024). "Future directions and emerging trends in medicinal plant DNA barcoding," in *Ethnopharmacology and OMICS advances in medicinal plants*, vol. 1. (Amsterdam, Netherlands: Elsevier), 219–237.
- Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., and Wingett, S. (2010). FastQC: A quality control tool for high-throughput sequence data. *Babraham Bioinf.*
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Daniell, H., Lin, C.-S., Yu, M., and Chang, W.-J. (2016). Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* 17, 134. doi: 10.1186/s13059-016-1004-2
- Dayal, R., Singh, A., Ojha, R. P., and Mishra, K. (2015). Possible therapeutic potential of *Helicteres isora* (L.) and its mechanism of action in diseases. *J. Med. Plants Stud.* 3, 95–100.
- Dobrogowski, J., Adamiec, M., and Luciński, R. (2020). The chloroplast genome: A review. *Acta Physiol. Plant* 42, 98. doi: 10.1007/s11738-020-03091-7
- Greiner, S., Lehwark, P., and Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47, W59–W64. doi: 10.1093/nar/gkz238
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., dePamphilis, C. W., Yi, T. S., et al. (2020). GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21, 241. doi: 10.1186/s13059-020-02154-5
- Kanthale, P. R., and Biradar, S. (2017). Pharmacognostic study of *Helicteres isora* L. *Pharm. Biol. Eval* 4, 47–51. doi: 10.26510/2394-0859.pbe.2017.07
- Kumbhani, N. R., Kuvad, R. P., and Thaker, V. S. (2017). Development of linear model for leaf area measurement of two medicinally important plants: *Helicteres isora* L. and *Vitex negundo* L. *Biol. Biotechnol. (India)* 5, 57–60. doi: 10.7324/JABB.2017.50310
- Kuznetsov, A., and Bollin, C. J. (2021). "NCBI Genome Workbench: Desktop software for comparative genomics, visualization, and GenBank data submission," in *Multiple sequence alignment: Methods and protocols*. Ed. K. Katoh (New York, NY, United States: Humana Press), 261–295. doi: 10.1007/978-1-0716-1036-7_16
- Lestari, R., Magandhi, M., Hariri, M. R., Noviady, I., Nugroho, A., and Indriani, F. (2024). Characterization of the complete chloroplast genome of the endangered and endemic Bornean fruit *Artocarpus tamaran* Becc. *Front. Plant Sci.* 15, 1513364. doi: 10.3389/fpls.2024.1513364
- Meena, R. S., Das, A., Yadav, G. S., and Lal, R. (2018). *Legumes for soil health and sustainable management* (Singapore: Springer).
- Nguyen, H. D., Vu, M. T., and Do, H. D. K. (2024). Characterization of the complete chloroplast genome of *Helicteres hirsuta* Lour. 1790 (Helicteriadeae: Malvaceae). *Mitochondrial DNA Part B* 9, 568–573. doi: 10.1080/23802359.2024.2345794
- Nugroho, A., Rustam, E., Widyani, N., Indriani, F., Sudrajat, D. J., Salim, M. A., et al. (2025). Complete assembly of the organellar genome of *Rubroshorea johorensis* utilizing advanced long-read sequencing technologies. *Front. Genet.* 16, 1574266. doi: 10.3389/fgene.2025.1574266
- Nuroniah, H. S., Nugroho, A., Indriani, F., Suita, E., Basyah, B., Susanna, S., et al. (2025). Chloroplast genome of *Myristica fragrans*: Assembly, annotation, and phylogenetic relationships. *For. Sci. Technol.* 21, 207–214. doi: 10.1080/21580103.2025.2482080
- Okonechnikov, K., Golosova, O., Fursov, M., and UGENE Team (2012). Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* 28, 1166–1167. doi: 10.1093/bioinformatics/bts091
- Olejniczak, S. A., Łojewska, E., Kowalczyk, T., and Sakowicz, T. (2016). Chloroplasts: State of research and practical applications of plastome sequencing. *Planta* 244, 517–527. doi: 10.1007/s00425-016-2531-7
- Ray, S., Satya, P., Sharma, L., Roy, S., Bera, A., Santra, S., et al. (2022). "Model plants in genomics," in *Plant genomics for sustainable agriculture* (Singapore: Springer), 241–264.
- Richards, C. L., Alonso, C., Becker, C., Bossdorf, O., Bucher, E., Colomé-Tatché, M., et al. (2017). Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecol. Lett.* 20, 1576–1590. doi: 10.1111/ele.12858
- Shi, L., Chen, H., Jiang, M., Wang, L., Wu, X., Huang, L., et al. (2019). CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47, W65–W73. doi: 10.1093/nar/gkz345
- Song, B. H., and Mitchell-Olds, T. (2011). Evolutionary and ecological genomics of non-model plants. *J. Syst. Evol.* 49, 17–24. doi: 10.1111/j.1759-6831.2010.00107.x