ARTICLE

Complete chloroplast genome of *Abutilon indicum* (L.) Sweet and comparative chloroplast genomics of Malvoideae

Crop Breeding and Applied Biotechnology
26(2): e54492621, 2026
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332026v26n2a16

Thi Thanh Nga Le[1] and Hoang Dang Khoa Do[1,2] ✉

Abstract: *Abutilon indicum (L.) Sweet, commonly known as Indian mallow, is a member of the family Malvaceae and possesses both ornamental and medicinal properties. Although this species has been utilized in folk medicine, the lack of genomic data has hindered further studies on A. indicum. Consequently, the current study sequenced and characterized the chloroplast genome of A. indicum. This genome is 160,066 bp in length, with a GC content of 36.99%, and exhibits a quadripartite structure comprising a large single-copy (88,733 bp), a small single-copy (20,037 bp), and two inverted repeats (25,648 bp each). Comparative analyses revealed conservation in genome structure, gene content, and repeat compositions among Abutilon chloroplast genomes. Additionally, six hypervariable regions were identified among the examined Malvoideae chloroplast genomes, including accD, matK-trnK-UUU, ndhA, ndhD, rps12-clpP1, and ycf1. Phylogenetic analysis indicated the paraphyletic status of the Abutilon genus.*

Keywords: *Indian mallow, Malvaceae, phylogenetics, plastid genome*

INTRODUCTION

Abutilon Mill. is a genus within the family Malvaceae and contains 176 accepted species distributed across tropical and subtropical regions (Plants of the World Online 2025). Previous studies have demonstrated that extracts from *Abutilon* species are rich in alkaloids, triterpenes, flavonoids, sterols, sphingolipids, and phenolic acids. These compounds exhibit antibacterial, antifungal, anti-inflammatory, anticancer, and antidiabetic properties (Li et al. 2024, Hassan et al. 2021). *Abutilon indicum* (L.) Sweet is a perennial shrub found in the Mascarenes, tropical and subtropical Asia, and the West Pacific (Plants of the World Online 2025). Similar to other *Abutilon* species, extracts of *A. indicum* contain various compounds of flavonoids, alkaloids, and terpenoids and possess antioxidant, antidiabetic, and antimicrobial activities (Musthafa et al. 2022, Sunil et al. 2023, Jadhav and Patil 2024). Although the phytochemical and pharmacological characteristics of *Abutilon* species have been extensively studied, genomic data (e.g., chloroplast, mitochondrial, and nuclear genomes) for the *Abutilon* genus remain limited. To date, only three out of the 176 *Abutilon* species have reported chloroplast genomes: *Abutilon megapotamicum*, *Abutilon fruticosum*, and *Abutilon theophrasti* (Lv et al. 2021, Alzahrani 2021, Maurya et al. 2025). Therefore, expanding the genomic data is essential for further studies examining the evolution, population genetics, and molecular identification of *Abutilon* species.

***Corresponding author:**
E-mail: dhdkhoa@ntt.edu.vn

Scientific Editor:
Luiz Antônio dos Santos Dias

Received: 4 November 2025
Accepted: 28 January 2026
Published: 20 February 2026

[1] Nguyen Tat Thanh University, NTT Hi-Tech Institute, Nguyen Tat Thanh Street, 754000, Ho Chi Minh City, Vietnam
[2] Nguyen Tat Thanh University Center for Hi-Tech Development, Saigon Hi-Tech Park, Vo Chi Cong Street, 721400, Ho Chi Minh City, Vietnam

The chloroplast genome (cpDNA) is a vital component of land plants responsible for photosynthesis (Dobrogojski et al. 2020). Due to its high conservation regarding genome structure and gene content, cpDNA has been widely employed to elucidate evolutionary history, phylogenetic relationships, and genomic variations (Zhu et al. 2016, Gitzendanner et al. 2018, Nguyen and Ho 2024). Various studies have examined cpDNAs among Malvaceae members, yielding insightful results for exploring evolution (Zhong et al. 2024, Jung et al. 2024, Maurya et al. 2025). Although records of *A. indicum* are available in the GenBank database under accession numbers PP897811 and OP581186, no published papers have reported on these available cpDNAs. Therefore, in the current study, we sequenced and characterized the complete cpDNA of *A. indicum* collected in Vietnam using next-generation sequencing technology. Additionally, comparative genomic analyses were conducted to assess nucleotide diversity, repeat content, and structural variation. Phylogenetic relationships between *A. indicum* and related Malvoideae taxa were reconstructed based on maximum likelihood and Bayesian inference methods using a dataset of 79 protein-coding genes (CDS). This study provides the first comprehensive characterization of *A. indicum* cpDNA and an updated comparative chloroplast genomic analysis of the Malvoideae subfamily, offering valuable resources for further studies examining evolution and comparative genomics in Malvoideae and Malvaceae.

MATERIAL AND METHODS

Plant sampling, DNA extraction, and sequencing

Healthy leaf samples of *A. indicum* were collected in Ho Chi Minh City (lat 10° 39' 21.8" N and long 106° 35' 46.9" E), Vietnam. The authors conducted species delimitation based on morphology, and the voucher specimen was stored at the Functional Genomics Research Center, NTT Hi-Tech Institute, Nguyen Tat Thanh University, under voucher number NTT_FGRC_20240912001 (contact person: Hoang Dang Khoa Do, dhdkhoa@ntt.edu.vn). The *A. indicum* leaf samples were dried using silica gel beads and stored under dry conditions at room temperature until further experiments. Total genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, USA) following the manufacturer's instructions. DNA quality was then assessed using gel electrophoresis (1% agarose gel) and a NanoDrop OneC Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific, USA). A high-quality DNA sample (concentration > 100 ng μL^{-1} and A_{260nm}/A_{230nm} ratio ~ 1.8 – 2.0) was used to prepare a sequencing library with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, UK). Subsequently, paired-end reads (150 bp) were generated using the Illumina NextSeq 550 platform (Illumina, USA) at KTest Science Co. Ltd (Ho Chi Minh City, Vietnam).

Genome assembly, annotation, and comparative analyses

All raw sequencing reads were qualified and filtered using the fastp v0.24.1 tool to obtain high-quality reads with quality scores > 20, a minimum length of 100 bp, and no ambiguous (N) bases (Chen et al. 2018). The filtered reads were then assembled into a complete cpDNA sequence using NOVOPlasty v4.3.5 (Dierckxsens et al. 2017), with *Abutilon megapotamicum* (NC_077649) as the reference. The complete cpDNA was annotated using the GeSeq program (Tillich et al. 2017). Transfer RNA sequences were verified using tRNAscan-SE, and the start and stop codons of CDSs were confirmed using Geneious Prime 2024.0.1 (Chan and Lowe 2019). The cpDNA was illustrated using OGDRAW (Greiner et al. 2019).

To assess the repeat content of *A. indicum* cpDNA and compare it with the three available *Abutilon* cpDNAs in the GenBank database, including *Abutilon megapotamicum*, *Abutilon theophrasti*, and *Abutilon fruticosum*, Phobos (Mayer 2006-2010) was used to identify various types of short sequence repeats, including mono-, di-, tri-, tetra-, penta-, and hexanucleotides, with parameters of 10, 6, 4, 3, 3, and 3, respectively. Four types of long repeats (e.g., forward, reverse, palindromic, and complementary) were identified using REPuter with a minimum length of 20 bp and a Hamming distance of three (Kurtz et al. 2001). The junctions among large single-copy (LSC), small single-copy (SSC), and inverted repeat (IR) regions in the complete cpDNAs of four *Abutilon* species and 18 other taxa, representing 18 genera of Malvoideae, were located and compared using IRplus with default settings (Menéndez et al. 2023). Additionally, to evaluate nucleotide diversity, all available complete cpDNAs of Malvoideae were searched and retrieved from the GenBank database. Consequently, 27 complete cpDNAs representing 19 genera of the Malvoideae subfamily were downloaded and combined with that of *A. indicum* to create a dataset, which was aligned using MUSCLE v5 with default settings (Edgar 2022). The aligned sequences were then used to evaluate nucleotide diversity using DnaSP6 with a window size

of 1000 and a step of 200 (Rozas et al. 2017). Synonymous (Ks) and nonsynonymous (Ka) substitutions of 79 CDSs of *Abutilon* cpDNAs were also estimated using DnaSP6, with Ka/Ks ratios indicating purifying or positive selection if Ka/Ks < 1 and Ka/Ks > 1, respectively.

Phylogenetic analysis

The available complete cpDNAs of 29 species, representing 19 genera within the Malvoideae subfamily, and two outgroups, including *Coelostegia griffithii* (NC_086661) from the Helicteroideae subfamily and *Theobroma cacao* (NC_014676) from the Byttnerioideae subfamily, were retrieved from the NCBI database to reconstruct a phylogenetic tree. Previously, chloroplast CDSs have proven effective in resolving phylogenetic relationships among flowering plants (Li et al. 2021). Therefore, in the current study, 79 CDSs from *A. indicum* and other downloaded cpDNAs were extracted and aligned using MAFFT v7 (Katoh and Standley 2013). The aligned sequences were used to find the optimal evolutionary model using jModeltest2, which identified the GTR+I+G model as the best fit (Darriba et al. 2012). The phylogenetic tree was reconstructed with the best substitution model using the Maximum Likelihood (ML) method with ultrafast bootstrap analysis and 1000 bootstrap replications, and the Bayesian Inference (BI) method with 1,000,000 generations. The ML phylogenetic tree was generated using the IQ-TREE web server (http://iqtree.cibiv.univie.ac.at), while BI was performed using MrBayes v3.2.7a (Huelsenbeck and Ronquist 2001). Finally, the phylogenetic tree was illustrated using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

RESULTS AND DISCUSSION

Features of the *Abutilon indicum* chloroplast genome

A total of 383,482 out of 37,778,212 reads were assembled, resulting in an average depth of 339x (Figure S1) to obtain a complete *A. indicum* cpDNA of 160,066 bp in length, with a GC content of 36.99% (Figure 1). This genome consists of four typical regions: an LSC, an SSC, and two IR regions, measuring 88,733 bp, 20,037 bp, and 25,648 bp in length, respectively. Additionally, the cpDNA contains 113 unique genes, including 79 CDSs, four rRNA genes, and 30 tRNA genes (Table S1). Most of the CDSs do not contain introns, except for nine genes (e.g., *atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl2*, *rpl16*, *rpoC1*, and *rps16*) that each contain one intron each and two genes (*clpP1* and *pafI*) that contain two introns (Figure S2). The *rps12* gene is transpliced, with exon 1 located in the LSC region, and exons 2 and 3 situated in the IR regions (Figure S2). The complete cpDNA of *A. indicum* is similar to those of angiosperms regarding genome structure and gene content (Dobrogojski et al. 2020).

Among published Malvaceae cpDNAs, the stability of the quadripartite structure and gene content was also observed (Zhong et al. 2024, Jung et al. 2024, Maurya et al. 2025). However, the loss of *clpP1* and *rpl32* in the cpDNA of *Durio zibethinus* (Helicteroideae), the loss of *rpl2* in *Talipariti hamabo* (Malvoideae) and *Ceiba pentandra* (Bambacoideae), and the pseudogenization of the chloroplast *infA* gene in various malvaceous species have been reported (Zhong et al. 2024, Ruang-areerate et al. 2025). Within Malvoideae cpDNAs, pseudogenization and complete coding sequences of *infA* were found in different species (Ruang-areerate et al. 2025). Similarly, the loss of *rpl2* and *rpl32* has not always been observed in *Durio zibethinus* cultivars (Shearman et al. 2020, Huy et al. 2024). These findings suggest a high degree of variation in this gene during speciation within Malvaceae. Although no gene loss has been reported in Malvoideae cpDNAs, the large number of species and their widespread distribution may conceal different genomic events, such as gene duplication, deletion, and inversion, which require further investigations.

Among the examined *Abutilon* cpDNAs, the *atpF*, *ccsA*, *ndhD*, and *rpoC2* genes exhibited positive selection (Ka/Ks > 1), while all the remaining genes displayed negative selection (Figure S3). In previous studies of *Hibiscus* cpDNAs, the *psbH* and *rpl22* genes showed positive selection, whereas *atpF*, *ccsA*, *ndhD*, and *rpoC2* exhibited negative selection (Zhu et al. 2025). Across Malvaceae cpDNAs, the *clpP1*, *psbI*, *psbT*, *rps7*, *rpl22*, and *rpl23* genes showed Ka/Ks values > 1, indicating positive selection (Abdullah et al. 2020). These findings indicate varying degrees of evolutionary pressure among genera of Malvaceae, necessitating additional malvaceous samples to clarify evolutionary patterns within the family.

Analysis of small single repeats (SSRs) in four *Abutilon* cpDNAs identified six SSR types; however, hexanucleotide repeats were only found in *A. megapotamicum* (Figure 2A). Among the six SSR types, mononucleotide repeats were the

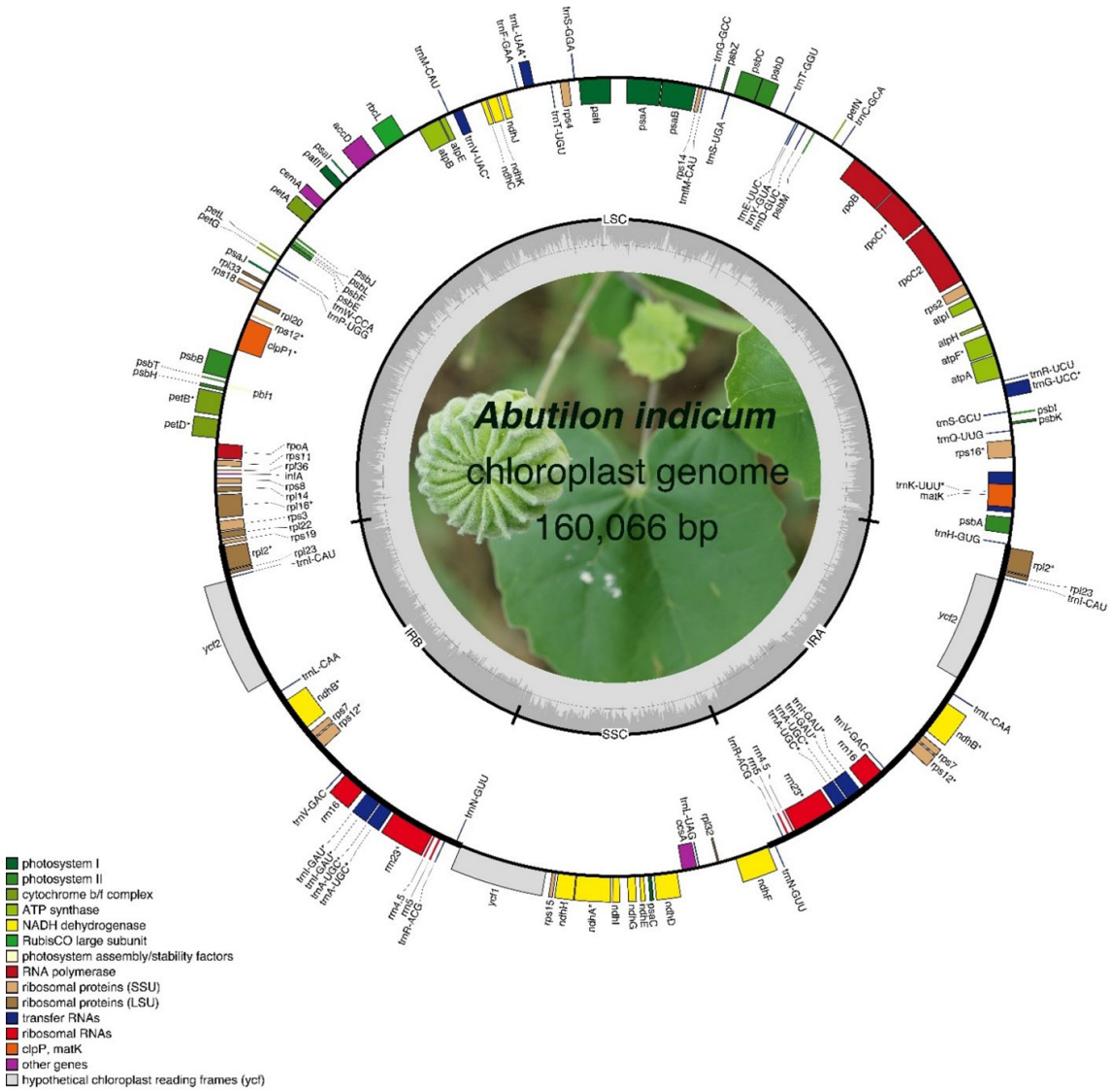

Figure 1. Chloroplast genome map of *A. indicum.* The grey inner circle illustrates the location of LSC, SSC, and IR regions with the CG contents. The outer circle indicates the genes, color-coded according to their functional classification. While the genes outside the circle follow anticlockwise transcription directions, those inside follow clockwise ones. The gene functional classification is shown in the bottom corner. The inner photo is the fruit of *Abutilon indicum*, photographed by Thi Thanh Nga Le.

most abundant, followed by tetranucleotide repeats. The cpDNA of *A. theophrasti* had the largest number of SSRs (85 repeats), followed by *A. fruticosum* (79 repeats), *A. indicum* (78 repeats), and *A. megapotamicum* (69 repeats). Further examination of the SSRs in *A. indicum* cpDNA indicated a predominance of mononucleotide repeats, which accounted for 76.9% of all repeats, followed by tetranucleotide repeats (12.8%) and pentanucleotide repeats (5.1%) (Figure 2B). Additionally, the AT motif was abundant, while the GC motif was limited to tetranucleotide and pentanucleotide repeats in the SSRs of *A. indicum* cpDNA. The long repeat analysis revealed a similar trend across the four examined *Abutilon* cpDNAs (Figure 2C). Specifically, forward repeats were the most common type, followed by palindromic and reverse repeats, whereas complementary repeats were rare and detected only in *A. megapotamicum* cpDNA. Furthermore, the quantity of long repeats was identical (50 repeats), although the repeat components varied among the four examined *Abutilon* cpDNAs.

Previous studies have reported an abundance of mononucleotide and dinucleotide SSRs in the cpDNAs of Malvaceae (Alzahrani 2021, Zhong et al. 2024, Maurya et al. 2025). At the generic level, cpDNAs of *Abelmoschus* and *Abutilon* (this study) also exhibited a similar trend, with a majority of mononucleotide and dinucleotide types among observed SSRs (Figure 2) (Li et al. 2020). Like SSRs, long repeats in published Malvaceae cpDNAs were rich in forward and palindromic types, while reverse and complementary types were rare among subfamilies and genera (Abdullah et al. 2020, Li et al. 2020, Zhong et al. 2024, Maurya et al. 2025). A previous study explored shared SSRs and long repeats among cpDNAs of 145 species belonging to 42 genera across nine subfamilies of Malvaceae (Wu et al. 2023). Although only 145 out of 4225 Malvaceae species were examined, the shared SSRs and long repeats indicated a conserved pattern of repeat content among Malvaceae species, providing valuable information for further population genetic analysis.

Nucleotide diversity and LSC/IR/SSC junctions among Malvoideae chloroplast genomes

Exploration of nucleotide diversity across 28 Malvoideae cpDNAs identified six hypervariable regions with Pi values ≥ 0.04, including *accD* (0.0449), *matK-trnK (UUU)* (0.04162), *ndhD* (0.04142), *ndhA* (0.04266), *rps12-clpP1* (0.05621), and *ycf1* (0.06582) (Figure 3). Most of these hypervariable regions were located in the LSC and SSC regions, with the *ycf1* gene exhibiting the highest Pi value, followed by the *rps12-clpP1* region. In contrast, the IR regions exhibited lower Pi values (≤ 0.02) compared to the LSC and SSC regions. Previous comparative genomic studies have identified various hypervariable regions in the cpDNAs of Malvaceae (Wang et al. 2021, Zhong et al. 2024, Maurya et al. 2025). For example, six hypervariable regions, including *ndhF-rpl32*, *petA-pabJ*, *psbC-trnS-UGA*, *rbcL-accD*, *rpl32-trnL-UAG*, *rps16-trnQ-UUG*, and *ycf1*, were found in six Malvaceae cpDNAs (Zhong et al. 2024). Another study examining 31 Malvaceae cpDNAs revealed *clpP1*, *matK*, *ndhD-ccsA*, *ndhF*, *rpl22*, *rpl22-rps19*, *rpl32-ndhF*, *rps19-rpl2*, *trnH-GUG-psbA*, *trnL-UAG-rpl32*, and *ycf1* as highly variable regions (Wang et al. 2021). Within the genera of Malvaceae, cpDNAs exhibited various variable regions, such as *ccsA-ndhD*, *petD-rpoA*, *petL*, *psbZ-trnG-GCC*, *trnH-GUG-psbA*, *trnN-GUU-ndhF*, *trnT-GGU-psbD*, and *ycf1* in *Sida* species and *accD-psaI*, *ndhF-ycf1*, *petA-psbJ*, *psbZ-trnG-GCC*, and *trnN-GUU-ndhF* in *Malva* species (Maurya et al. 2025). Although different variable regions have been identified across malvaceous cpDNAs, *ycf1* is commonly found among subfamilies and within genera, suggesting *its utility* as a sequence for exploring phylogeny and mining molecular markers in Malvaceae.

Comparative analysis revealed different features of gene content in the junctions among the LSC, SSC, and IR regions of the examined Malvoideae cpDNAs (Figure S4). Specifically, the LSC/IRb junction (JLB) is commonly located within the *rps19* genes (ranging from 2 bp to 14 bp toward

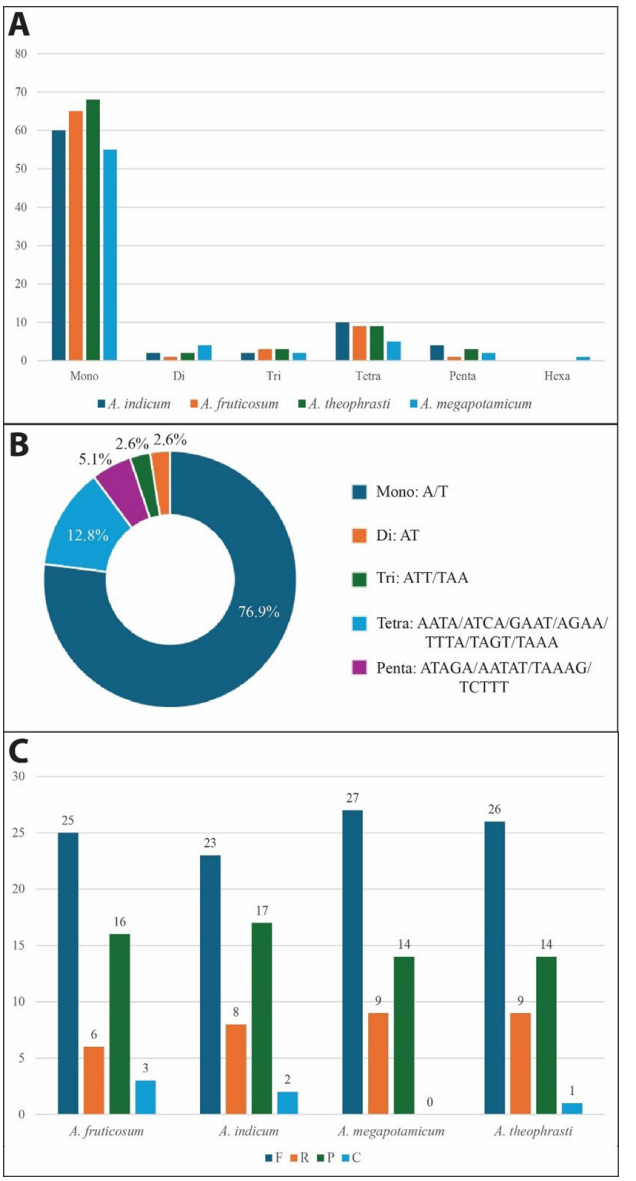

Figure 2. Repeat content of *Abutilon* chloroplast genomes. A. The SSR contents of four *Abutilon* species; B. The AT content of SSR in the *Abutilon indicum* chloroplast genome; C. Long repeat of four *Abutilon* species. Four letters, including F, R, P, and C, represent four types of long repeats: forward, reverse, palindromic, and complementary, respectively. The y-axis represents the number of repeats.

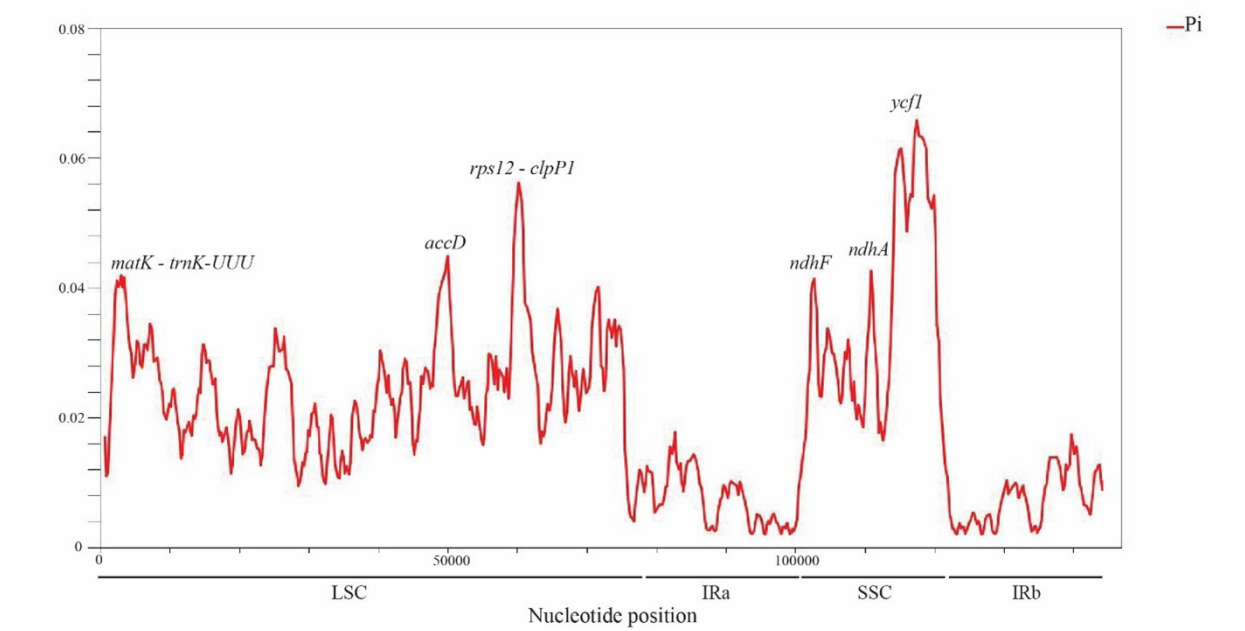

Figure 3. Nucleotide diversity of 28 Malvoideae chloroplast genomes. The six regions with the highest Pi values are labeled. LSC indicates the large single-copy region; IR denotes the inverted repeat region; SSC indicates the small single-copy region.

the start codon). However, in *Hibiscus mutabilis* and *Malvaviscus penduliflorus*, the JLB occurred within the intergenic spacer (IGS) between *rpl2* and *rps19* genes. A notable case was observed in *Abelmoschus esculentus*, in which the JLB was located within the *rpl16* intron. Consequently, the gene content of the LSC/IRa junction (JLA) was similar to that of the JLB regarding gene composition in the IRa region. Nevertheless, the LSC site of JLA witnessed the movement of *trnH-GUG* to the junction site. Most examined Malvoideae cpDNAs had a space ranging from 1 bp to 332 bp between the *trnH-GUG* and the junction sites. Meanwhile, *Gossypium populifolium*, *Kosteletzkya pentacarpos*, *Thespesia populnea*, and *Talipariti hamabo* showed adjacency of the *trnH-GUG* and the junction site. Similar to the special JLB, the JLA of *A. esculentus* was located within *trnH-GUG*, which was not present in other examined species.

In contrast to the variety of LSC and IR junctions, the boundary between the SSC and IR regions (JSA and JSB) was commonly found in the *ycf1* gene, varying from 3 bp to 1131 bp toward the start codon. However, the JSA and JSB of *Alcea rosea*, *Althaea officinalis*, *Malva parviflora*, *Malvastrum coromandelianum*, *Sidalcea hendersonii*, and *Sida acuta* were located in the IGS between *ndhF* and *trnN-GUU* and between *trnN-GUU* and *ycf1*, respectively (Figure S4).

The specific content of the IR regions has been previously demonstrated among angiosperms, gymnosperms, ferns, and green algae, which exhibited an expansion of the IR region in the cpDNAs of angiosperms (Zhu et al. 2016). The current study reveals that the IR junctions of Malvoideae cpDNAs are similar to those observed in other angiosperms, except for a notable expansion of the IR region to *rpl16* and *trnH-GUG* in *A. esculentus* (Figure S4). Further observations in the IR junctions of *Abelmoschus* indicate a conservation of the IR junction content, suggesting specific variation within *Abelmoschus* that may be useful for mining molecular markers (Li et al. 2020). Additionally, the presence of two types of IR/SSC junctions in Malvoideae implies different evolutionary scenarios during the speciation of Malvoideae, which require more genomic data in future studies.

Phylogenetic relationships between *Abutilon indicum* and related Malvoideae taxa

Phylogenetic analysis based on 79 CDSs revealed the monophyly of Malvoideae, which consists of two main clades: clade C-I of Hibisceae and clade C-II of Gossypieae (C-IIa) and Malveae (C-IIb) (Figure 4). Although the monophyly of tribes within Malvoideae was reconstructed with high support values, non-monophyletic status was observed in the

Hibiscus, *Sida*, and *Abutilon* genera. Within the *Abutilon* genus, the target *A. indicum* was closely related to *A. fruticosum*, while other *A. indicum* samples formed a clade with *A. theophrasti*. Furthermore, *A. megapotamicum* formed a clade with *Callianthe picta*, which was sister to the remaining *Abutilon* species.

Previously, complete cpDNAs and CDS datasets exhibited incongruent phylogenetic relationships among Malvaceae subfamilies (Wang et al. 2021). However, both datasets have consistently yielded high support values for relationships

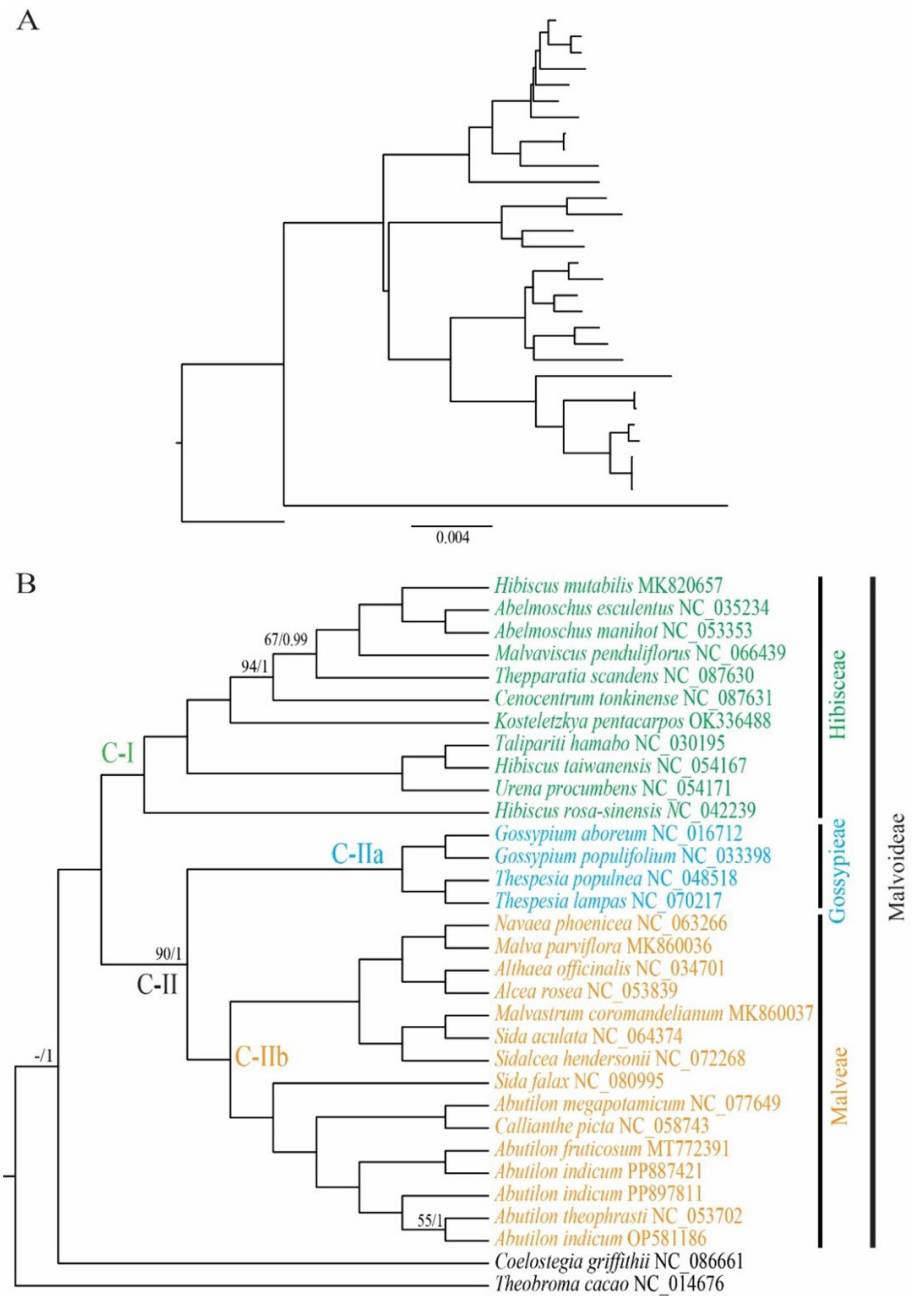

Figure 4. Phylogenetic tree of 30 Malvoideae species and related taxa based on 79 protein-coding genes. A. The phylogram format of the phylogenetic tree. The scale bar indicates the number of nucleotide substitutions per site. B. The cladogram format of the phylogenetic tree. Only bootstrap support values < 100% and posterior probabilities < 1 are shown. The bold name indicates the target species of the present study. Dashes indicate no value.

among malvaceous genera. Additionally, CDSs have been successfully employed to reconstruct relationships among plant families (Li et al. 2021). Therefore, in the current study, we reconstructed the phylogenetic relationships of Malvoideae species based on 78 CDSs, revealing a close relationship between Gossypieae and Malveae. This relationship was also confirmed by another previous study of Malvaceae inferred from 78 chloroplast CDSs (Wang et al. 2021). However, another phylogenetic study based on 300 low-copy nuclear genes supported the close relationship between Gossypieae and Malveae (Colli-Silva et al. 2025). Together, these findings indicate incongruence between the cpDNA and nuclear genome-based phylogenetic analyses within Malvoideae, necessitating further investigation for a clearer understanding of relationships among Malvaceae members.

The non-monophyletic status of *Hibiscus* and *Sida* has been previously reported based on cpDNAs, which was also found in the current study (Hanes et al. 2024, Maurya et al. 2025). Resolving the phylogenetic relationships of these genera requires further investigation based on both molecular data (e.g., nuclear, mitochondrial, and chloroplast genomes) and morphological characteristics of *Hibiscus* and *Sida*. In the case of *Callianthe*, which was previously recognized as a new genus distinct from *Abutilon* and *Bakeridesia*, the monophyly of *Callianthe* species was reconstructed based on the ITS region (including ITS1, the *5.8S* subunit, and ITS2) and morphological characteristics (Donnell et al. 2012). However, the placement of *Callianthe picta* within *Abutilon* species and its close relationship to *A. megapotamicum* based on cpDNAs indicate ongoing taxonomic issues with these two genera (Figure 4). In general, the incongruences of phylogenetic relationships inferred from different datasets highlight the complex evolution within Malvaceae and necessitate more comprehensive analyses.

The newly completed cpDNA of *A. indicum* expands the genomic data for the *Abutilon* genus. However, only one sample of *A. indicum* collected in Vietnam was used in the current study. Therefore, additional samples of *A. indicum* from its widespread distribution in the East and West Himalayas, Northern and Western Australia, South Asia, Southwest Asia, and Southeast Asia should be collected and analyzed to trace genomic evolution. Furthermore, the availability of four *Abutilon* cpDNAs remains limited compared to the 178 accepted *Abutilon* species. Thus, more genomic data, including organelle and nuclear genomes, should be explored to elucidate the evolutionary history of the *Abutilon* genus. Additionally, the lack of genomic data in the tribe Kydieae hinders the exploration of phylogenetic relationships within Malvoideae. Therefore, future studies must prioritize the sequencing and characterization of the cpDNAs of tribe Kydieae members, such as *Kydia*.

ACKNOWLEDGMENTS

We acknowledge Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam, for supporting this study, particularly through the provision of high-performance computing devices. The manuscript includes supplementary files that are available from the corresponding author upon request.

CREDIT STATEMENT

All authors contributed to the study's conception and design. TTN Le and HDK Do prepared the material, collected, and analyzed the data. The initial draft was written by TTN Le, and all authors provided feedback on the following versions. All authors read and approved the final manuscript.

DATA AVAILABILITY

The datasets generated and/or analyzed during the current research are available from the corresponding author upon reasonable request.

REFERENCES

Abdullah, Mehmood F, Shahzadi I, Waseem S, Mirza B, Ahmed I and Waheed MT (2020) Chloroplast genome of *Hibiscus rosa-sinensis* (Malvaceae): comparative analyses and identification of mutational hotspots. **Genomics 112**: 581-591.

Alzahrani DA (2021) Complete chloroplast genome of *Abutilon fruticosum*: genome structure, comparative and phylogenetic analysis. **Plants 10**: 270.

Chan PP and Lowe TM (2019) tRNAscan-SE: Searching for tRNA genes in genomic sequences. **Methods Molecular Biology 1962**: 1-14.

Chen S, Zhou Y, Chen Y and Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. **Bioinformatics 34**: i884-i890.

Colli-Silva M, Pérez-Escobar OA, Ferreira CDM, Costa MTR, Gerace S, Coutinho TS, Yoshikawa VN, Antonio-Domingues H, Hernández-Gutiérrez R, Bovini MG, Duarte MC, Cheek M, Chase MW, Fay MF, Christenhusz MJM, Dorr LJ, Schoeplein C, Corcoran M, Roy S, Cable S, McLay T, Maurin O, Forest F, Baker WJ and Antonelli A (2025) Taxonomy in the light of incongruence: an updated classification of Malvales and Malvaceae based on phylogenomic data. **Taxon 74**: 361-385.

Darriba D, Taboada GL, Doallo R and Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. **Nature Methods 9**: 772.

Dierckxsens N, Mardulyn P and Smits G (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. **Nucleic Acids Research 45**: e18.

Dobrogojski J, Adamiec M and Luciński R (2020) The chloroplast genome: a review. **Acta Physiologiae Plantarum 42**: 98.

Donnell AA, Ballard HE and Cantino PD (2012) Callianthe (Malvaceae): a new genus of Neotropical Malveae. **Systematic Botany 37**: 712-722.

Edgar RC (2022) Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. **Nature Communications 13**: 6968.

Gitzendanner MA, Soltis PS, Wong GK-S, Ruhfel BR and Soltis DE (2018) Plastid phylogenomic analysis of green plants: a billion years of evolutionary history. **American Journal of Botany 105**: 291-301.

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 Research 47**: W59-W64.

Hanes MM, Blanchard OJ, Valencia DJ, McLay T, Abbott JR, McDaniel SF, Barrett RL, Mathews S and Neubig KM (2024) Phylogenetic relationships within tribe Hibisceae (Malvaceae) reveal complex patterns of polyphyly in *Hibiscus* and *Pavonia*. **Systematic Botany 49**: 77-116.

Hassan M, Haq SM, Rasool A, Fatima S, Ashraf A, Zulfajri M and Hanafiah MM (2021) Ethnobotanical properties and traditional uses of medicinal plant *Abutilon theophrasti* Medik. In Tariq A and Khalid Rehman H (eds) **Medicinal and aromatic plants**. Springer International Publishing, Cham, Switzerland, p. 271-285.

Huelsenbeck JP and Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. **Bioinformatics 17**: 754-755.

Huy TG, Thi NPA, Do HDK and Khang DT (2024) The complete chloroplast genome of *Durio zibethinus* L. cultivar Ri6 (Helicteroideae, Malvaceae). **Mitochondrial DNA Part B 9**: 625-630.

Jadhav NR and Patil RB (2024) Analyzing the phytochemistry and bioactive compounds of *Abutilon indicum* for antidiabetic activity using TLC, HPTLC, GC-MS in aqueous, ethanol, and chloroform solvents for pharmacognosy determinations. **Journal of Pharmaceutical Research 23**: 40-48.

Jung J, Deng T, Kim YG, Kim C, Sun H and Kim JH (2024) Comparative phylogenomic study of East Asian endemic genus, *Corchoropsis* Siebold & Zucc. (Malvaceae s.l.), based on complete plastome sequences. **BMC Genomics 25**: 854.

Katoh K and Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and usability. **Molecular Biology and Evolution 30**: 772-780.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J and Giegerich R (2001) REPuter: the manifold applications of repeat analysis on a genomic scale. **Nucleic Acids Research 29**: 4633-4642.

Li HT, Luo Y, Gan L, Ma PF, Gao LM, Yang JB, Cai J, Gitzendanner MA, Fritsch PW, Zhang T, Jin JJ, Zeng CX, Wang H, Yu WB, Zhang R, van der Bank M, Olmstead RG, Hollingsworth PM, Chase MW, Soltis DE, Soltis PS, Yi TS and Li DZ (2021) Plastid phylogenomic insights into relationships of all flowering plant families. **BMC Biology 19**: 232.

Li J, Ye GY, Liu HL and Wang ZH (2020) Complete chloroplast genomes of three important species, *Abelmoschus moschatus*, *A. manihot* and *A. sagittifolius*: genome structures, mutational hotspots, comparative and phylogenetic analysis in Malvaceae. **PLoS One 15**: e0242591.

Li L, Pang L, Chen F, Zhang L, Zhou F, Li R and Yan H (2024) Abutilon theophrasti Medic.: a review of its traditional uses, phytochemistry, pharmacology, and pharmacokinetics. **Natural Product Communications 19**: 1934578X241257380.

Lv Y, Yi Y, Han R, Tong X and Takahashi H (2021) The complete chloroplast genome of *Abutilon theophrasti* medic (Malvaceae). **Mitochondrial DNA Part B 6**: 912-913.

Maurya S, Sukhramani G, Lee C, Jo S, Kim J, Jeong EJ, Zamora NA, Garcia R, Zhang Z, Choi S, Choudhary RK and Kim S-Y (2025) Comparative plastome evaluation, phylogenomic analysis, and DNA signatures of medicinally important malvaceous genera and their adulterants. **Scientific Reports 15**: 24285.

Mayer C (2006-2010) Phobos 3.3.12 Available at http://www.rub.de/ecoevo/cm/cm_phobos.htm. Accessed on September 1, 2025.

Menéndez CD, Poczai P, Williams B, Myllys L and Amiryousefi A (2023) IRplus: an augmented tool to detect inverted repeats in plastid genomes. **Genome Biology and Evolution 15**: evad177.

Musthafa SA, Dabdoub W, Sadiq M and Munuswamy-Ramanujam G (2022) Syringaldehyde isolated from *Abutilon indicum* Linn. leaves exhibits broad spectrum anti-microbial activity **Materials Today: Proceedings 50**: 335-339.

Nguyen MP and Ho VT (2024) Characterization of the complete chloroplast genome of *Camellia tienii* (Theaceae) and its relatives. **Crop Breeding and Applied Biotechnology 24**: e486824415.

Plants of the World Online (2025) Abutilon Mill. Available at <https://powo.science.kew.org/taxon/ urn:lsid:ipni.org:names:30000557-2>. Accessed on October 1, 2025.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE and Sánchez-Gracia A (2017) DnaSP 6: DNA

sequence polymorphism analysis of large data sets. **Molecular Biology and Evolution 34**: 3299-3302.

Ruang-areerate P, Sangsrakru D, Yoocha T, Kongkachana W, U-Thoomporn S, Prathip Na Thalang O, Chumriang P, Wanthongchai P, Tangphatsornruang S and Pootakham W (2025) Comparative chloroplast genomes and phylogenetic relationships of true mangrove species *Brownlowia* tersa and *Brownlowia argentata* (Malvaceae). **Current Issues in Molecular Biology 47**: 74.

Shearman JR, Sonthirod C, Naktang C, Sangsrakru D, Yoocha T, Chatbanyong R, Vorakuldumrongchai S, Chusri O, Tangphatsornruang S and Pootakham W (2020) Assembly of the durian chloroplast genome using long PacBio reads. **Scientific Reports 10**: 15980.

Sunil M, Vedavijaya T, Sree PK and Sayana SB (2023) Phytochemical analysis and antioxidant evaluation of the ethanolic extract of the leaves of *Abutilon indicum*. **Cureus 15**: e47703.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R and Greiner S (2017) GeSeq - versatile and accurate annotation of organelle genomes. **Nucleic Acids Research 45**: W6-W11.

Wang JH, Moore MJ, Wang H, Zhu ZX and Wang HF (2021) Plastome evolution and phylogenetic relationships among Malvaceae subfamilies. **Gene 765**: 145103.

Wu M, He L, Ma G, Zhang K, Yang H and Yang X (2023) The complete chloroplast genome of *Diplodiscus trichospermus* and phylogenetic position of Brownlowioideae within Malvaceae. **BMC Genomics 24**: 571.

Zhong Y, Bai B, Sun Y, Wen K, Qiao Y, Guo L, Deng H, Ye Y, Feng L and Feng X (2024) Comparative genomics and phylogenetic analysis of six Malvaceae species based on chloroplast genomes. **BMC Plant Biology 24**: 1245.

Zhu A, Guo W, Gupta S, Fan W and Mower JP (2016) Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. **New Phytologist 209**: 1747-1756.

Zhu Z, Liu J, Tang S, Ji Q, An X, Dong J, Luo X, Chen C, Liu T, Zou L, Li S, Luan M and An X (2025) Sequencing, assembly, and comparative evolutionary analysis of the chloroplast genome of Kenaf (*Hibiscus cannabinus* L.). **Genes (Basel) 16**: 1519.