Dhaka Univ. J. Biol. Sci. 34(1): 119-143, 2025 (January) DOI: https://doi.org/10.3329/dujbs.v34i1.79865

COMPARATIVE GENOMICS AND PHYLOGENETIC ANALYSIS OF COMPLETE CHLOROPLAST GENOME OF *SCAPHIUM SCAPHIGERUM* (WALL. *EX* G. DON) G. PLANCH

Sheikh Sunzid Ahmed And M. Oliur Rahman*

Department of Botany, Faculty of Biological Sciences, University of Dhaka, Dhaka 1000, Bangladesh

Keywords: chloroplast genome, molecular dating, phylogenetic analysis, *Scaphium scaphigerum*, Sterculioideae

Abstract

Scaphium scaphigerum (Wall. *ex* G. Don) G. Planch. is a threatened medicinal tree of Bangladesh, belonging to the family Malvaceae. This study unveiled the first chloroplast genome of the species, spanning 160,927 bp with a GC content of 36.89%. The plastome exhibited the typical orbicular quadripartite structure of angiosperms, consisting of a pair of inverted repeat regions (25502 bp), isolated by a large-single copy region (90012 bp) and a small single-copy region (19911 bp). The plastome harbored 128 genes in total, including 85 protein-coding genes, 35 tRNAs and 8 rRNAs. Comparative analyses with several other species revealed high synteny and lack of major rearrangements, along with similar gene order and GC content. The plastome contained 115 SSRs, predominantly comprising mononucleotides (78). Among the longer repeats, palindromic sequences were the most frequent (22). Nucleotide diversity analysis identified two hypervariable sites (*ycf1* and *ndhI*) in the small single-copy region, which will facilitate DNA barcoding endeavors. Phylogenetic analysis showed close alliance of *S. scaphigerum* within Sterculioideae, with strong bootstrap support. Molecular dating suggested that the species originated during the Lutetian age (48.23 MYA) of the Cenozoic era.

Introduction

Scaphium scaphigerum (Wall. *ex* G. Don) G. Planch. (Malvaceae), a threatened medicinal tree of Bangladesh thrives in evergreen forests, particularly in limestone-rich soils, where it plays a crucial role in the ecosystem(1). This species is characterized by its large, deciduous stature, growing up to 35 meters tall, often with prominent buttresses extending 1–2 meters above the ground. The bark of the tree is distinctive, with an outer grey-green to brown cracked surface and a fibrous, reddish inner layer. The leaves are 8-34 cm long and 6-17 cm width, sub-leathery to leathery, occasionally papery, and vary in shape from elliptic-oblong to ovate, with an orbicular to acute base and an acute apex. The inflorescences are densely

* Author for Correspondence: oliur.bot@du.ac.bd

stellate hairy, bearing yellow flowers with a purple base. The fruit comprises 2–5 papery follicles, each containing a single seed(1,2).

Taxonomic identification of *S. scaphigerum* is very important for the sustainable maintenance and preservation of this threatened therapeutic plant. Accurate identification ensures proper recognition of the species, which is vital for protecting its genetic resources and preventing its extinction(3). *S. scaphigerum* holds significant medicinal and cultural value across various regions. In China, the flesh is utilized to treat gastrointestinal disorders, while in Cambodia, Thailand, and Malaysia, the gelatinous mass formed from the fruit's outer shell is consumed as a delicacy and used to treat diarrhea. Additionally, the fruit gel powder has been shown to inhibit glucose absorption, exhibit antioxidant properties, and aid in weight control(1,4,5). Precise taxonomic identification ensures quality control in medicinal preparations, and strengthens conservation efforts, enabling communities to continue benefiting from its therapeutic uses. Complementing this, the studying chloroplast (Cp) genome of *S. scaphigerum* offers valuable insights into its evolutionary biology and adaptive mechanisms(6).

Cp genomes are significantly more conserved in nucleotide sequence, composition, and structure compared to mitochondrial and nuclear genomes, making them highly valuable for molecular systematics studies. The number of plastomes deposited in the GenBank database is increasing continuously, driven by advancements in genomic technologies. One notable innovation is the utilization of publicly accessible next-generation sequencing (NGS) data to assemble chloroplast genomes, eliminating the need for new wet-lab experiments(7). This approach dramatically reduces the time, costs, and resources typically required for genome sequencing. By tapping into this expansive repository of genetic data, it is feasible to efficiently assemble complete chloroplast genomes with great accuracy(8). The availability of public data makes Cp genome assembly more scalable and accessible and promotes cooperation and replicability within the scientific community(9).

Plastomes typically exhibit a quadripartite structure with two inverted repeats (IRs) flanking the large single-copy (LSC) and small single-copy (SSC) regions, which are crucial for photosynthesis and the biosynthesis of key macromolecules(10). The plastome of *S. scaphigerum* may reveal structural variations, such as expansions and contractions, that can influence gene content and organization, offering insights into its evolutionary history and environmental adaptation. Studying plastome diversity can enhance phylogenetic understanding, track species evolution, and inform conservation strategies for conserving this threatened medicinal species(11,12). Furthermore, integrating taxonomic identification with Cp genome insights plays an important role to safeguard the genetic diversity and medicinal potential of *S. scaphigerum* for future generations.

To date, no comprehensive study has explored the complete Cp genome of *S. scaphigerum*, limiting our understanding of its plastome-wide systematic positions and its origin on the geological timescale using appropriate genetic markers. Therefore, we aim to assemble the first full length plastome of *S. scaphigerum* and assess its characteristic features through phylogenetic and molecular dating approaches. These outcomes will enhance current understanding of the genetic classification and evolutionary history of *S. scaphigerum*, contributing significantly to its conservation efforts and advancing research on its medicinal and ecological significance.

Materials and Methods

Next-generation sequencing data retrieval and quality assessments: The whole genome sequencing (WGS) data of *Scaphium scaphigerum* (synonym: *Scaphium wallichii*) was analyzed to construct the complete Cp genome. The sequencing was carried out by the Guangxi Botanical Garden of Medicinal Plants (GBGMP) in China from leaf samples of the species, and the data was made publicly available under the accession ID "SRR25033356". GBGMP employed the Illumina NovaSeq 6000 platform, generating 17,376,023 spots and 5.2 Gb of base with a total size of 1.7 Gb during this process. The library (CPD03) utilized a random selection strategy and paired-end layout, providing comprehensive data for further analysis. In the present investigation, we retrieved the NGS raw reads of *S. scaphigerum* deposited by GBGMP from the NCBI SRA database. The quality of the extracted data was assessed using FASTQC v0.12.1 to ensure high-quality reads(9).

Assembly, annotation and GenBank submission: The Illumina reads were assembled into the complete plastome via the GetOrganelle v.1.7.7.0 pipeline(13). Subsequently, coverage analysis was conducted with UGENE(14), and the plastome was annotated using the CPGAVAS2 server(15). The orbicular plastome was visualized through OGDraw(16). The complete plastome has been deposited to the GenBank database under accession number TPA: BK067803.

Repeats and codon usage assessments: Lengthy repeat structures were detected using the REPuter server, while SSRs (simple sequence repeats) were identified through the MISA-Web server(17,18). In REPuter, all matching directions were considered for the analysis of longer repeats. SSRs were analyzed using the default settings of the MISA-Web server. Codon usage was assessed by MEGA v.11(19).

RNA editing sites and GC skewness: RNA editing sites in the plastome were examined using the PREPACT 3.0 server(20). The BLASTx module was utilized to identify forward editing sites (C→U), with *Gossypium hirsutum* L. (Malvaceae) as the reference database and an e-value threshold of 0.001. For GC content skewness analysis, the assembled plastome was uploaded in FASTA format to the Proksee server(21). Following initial processing, GC content and GC skew analysis options were applied to produce the orbicular map.

IR contraction and expansion: The quadripartite junctions and associated genes in *S. scaphigerum* were analyzed using IRscope(22). The annotation file for *S. scaphigerum* was uploaded along with files from closely related taxa. After generating the plot, it was downloaded to examine the IRs.

Genome rearrangement and collinearity analysis: The assembled plastome of *S. scaphigerum* was subjected to comparative genomic analysis using the mVISTA tool(23). In addition, Mauve v.20150226 was used to identify gene order similarities with other taxa(24). For collinearity analysis, the plastome was examined using the Circoletto server(25).

Nucleotide diversity analysis: Nucleotide diversity analysis commenced with the alignment of chloroplast genome sequences utilizing the MAFFT online tool to ensure precise sequence alignment across the studied genomes(26). Subsequently, nucleotide diversity was assessed with DnaSP v.5, with a window size of 600 bp and a step size of 200 bp respectively, allowing for a detailed examination of nucleotide diversity across the genome(27).

Molecular phylogenetic and dating endeavor: Molecular phylogenetic analysis was performed using MEGA v.11, employing the Maximum-Likelihood (ML) method. Molecular dating was conducted with the RelTime-ML submodule(19). The analysis began with the loading of aligned chloroplast genome sequences, and divergence times were estimated by calibration node based on the data from the TimeTree server(28).

Results and Discussion

NGS reads quality evaluation

The FASTQC analysis of both forward and reverse reads for *S. scaphigerum* revealed consistently high quality across all parameters. The forward reads exhibited mean, median, lower quartile, upper quartile, and percentile score of 30 for each base position from 1 to 150, indicating no variability or degradation throughout the sequence. Similarly, the reverse reads demonstrated identical quality metrics, with all bases maintaining a consistent score of 30. Both read sets passed the per-base sequence quality checks with exceptional accuracy and reliability, ensuring robust and reliable data for downstream analyses. Given the uniform and stable quality scores across all positions in both forward and reverse reads, trimming was deemed unnecessary, further confirming the integrity of the raw sequencing data. These quality results align with previously published studies(9,10), supporting the use of these raw reads for assembling the *S. scaphigerum* plastome.

Plastome structure and contents

The plastome of *S. scaphigerum* spans a total length of 160,927 bp, comprising 90,012 bp in the LSC, 19,911 bp in the SSC, and 51,004 bp in the two IR regions (Fig. 1). The plastome exhibits the typical quadripartite structure characteristic of closely related species. A comparison with the plastomes of *Firmiana calcarean* and *F. hainanensis*, close relatives of *Scaphium*, reveals similar genome sizes of 161,263 bp and 161,031 bp, respectively. Both *Firmiana* species contain two IRs (25,549 bp and 25,521 bp), separated by LSC regions (90,141 bp and 89,968 bp) and SSC regions (20,024 bp and 20,021 bp)(29). The close similarity between these species further validates the accuracy of the *S. scaphigerum* plastome assembly, confirming the consistency of its structural arrangement within the Sterculioideae subfamily.

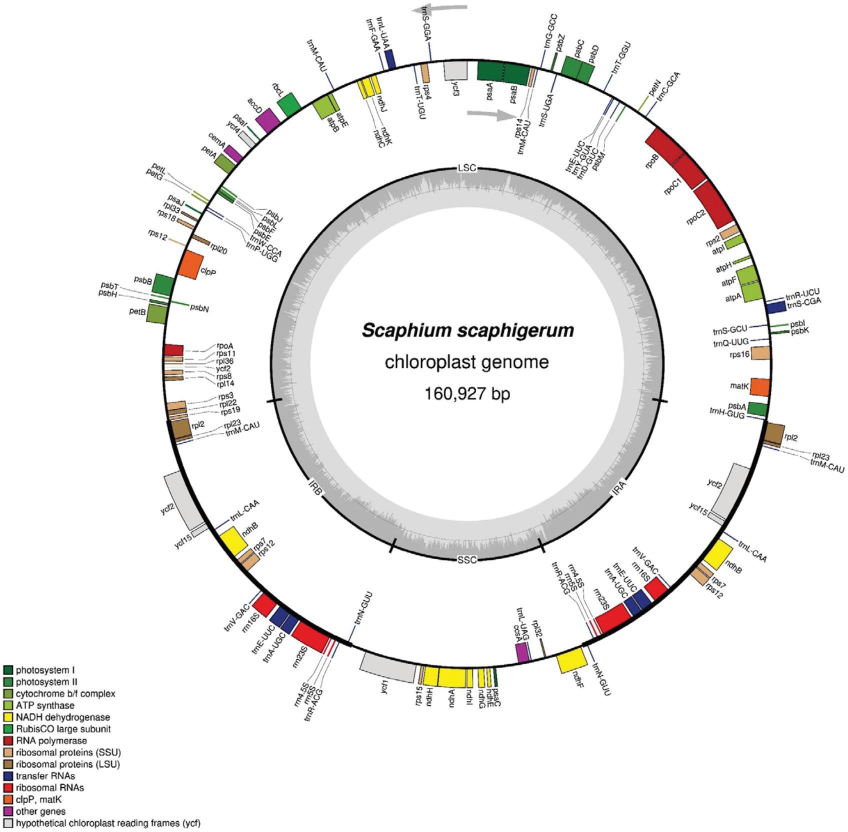


Fig. 1. Complete chloroplast genome of *S. scaphigerum* representing gene orders and quadripartite junction sites. Genes drawn within the circle are transcribed clockwise, while those drawn outside are transcribed counter clockwise. Genes are color-coded according to their functional groups. The inner circle represents GC contents throughout the plastome.

The comparative analysis of nucleotide composition across different compartments of the *S. scaphigerum* plastome revealed distinct patterns (Table 1). The overall plastome exhibited an AT content of 63.11% and a GC content of 36.89%. The LSC region had the highest AT content of 65.35%, with a correspondingly lower GC content of 34.65%, indicating a greater abundance of adenine and thymine bases. The SSC region further amplified this trend, with an AT content of 68.66% and a GC content of 31.34%. In contrast, the IRs showed a significantly higher GC content of 42.99%, with corresponding AT contents of 57.01%. These variations underscore the differences in nucleotide composition across the plastome compartments. The elevated AT ratio observed in the SSC and LSC regions compared to the IRs, suggests an evolutionary trend where these gene-rich regions favor AT-rich codons, potentially enhancing gene expression and protein function[30]. This higher AT content likely reflects increased recombination rates in these dynamic regions, promoting nucleotide variability. In contrast, the IR regions, which experience fewer recombination events, exhibit lower AT content and greater structural stability, preserving a more conserved nucleotide pattern. These differences underscore how distinct evolutionary pressures have shaped the plastome of *S. scaphigerum*, balancing functional demands with genomic stability, similar to patterns observed in other species[(31)].

Table 1. Nucleotide composition of the plastome of *S. scaphigerum*

Region	A (%)	T (U) (%)	C (%)	G (%)	C + G (%)	A + T (%)
cp Genome	31.13	31.98	18.61	18.28	36.89	63.11
LSC	31.85	33.49	17.80	16.85	34.65	65.35
SSC	34.64	34.03	14.87	16.47	31.34	68.66
IRA	28.54	28.47	22.29	20.70	42.99	57.01
IRB	28.47	28.54	20.70	22.29	42.99	57.01

Genome coverage analysis revealed an average depth of 990.165X across the complete plastome of *S. scaphigerum* (Fig. 2), with no regions showing zero coverage. This continuous and comprehensive coverage across the entire plastome ensures a highly accurate and complete assembly, eliminating concerns about missing or unsequenced regions that could compromise the structural integrity of the genome. The absence of zero-coverage areas is particularly crucial, as it confirms the reliability of the sequence data and reduces the risk of assembly gaps or erroneous base calls. This level of depth and consistency is critical for downstream analyses, such as precise gene annotation, evolutionary studies, and functional predictions, especially when examining regions with distinct nucleotide composition, such as the AT-rich LSC and SSC regions compared to the more stable IRs[(9)].

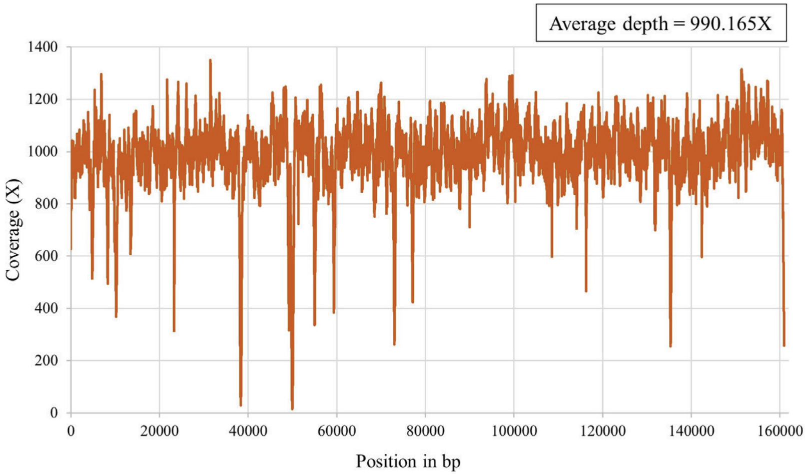


Fig. 2. Genome coverage analysis of the plastome of *S. scaphigerum.* No sequence gap was observed throughout the entire length, indicating reliable assembly of the Cp genome. The lowest coverage (14X) was noted at position 49989 and the highest coverage (1350X) was observed at position 31491.

Annotated genes of the Cp genome

S. scaphigerum plastome revealed a total of 128 genes, comprising 85 PCGs, 35 tRNAs, and eight rRNAs (Table 2). Consistent with other plastomes, genes involved in photosynthesis were well-represented, including five genes for photosystem I and 15 for photosystem II. The ATP synthase complex was encoded by six genes, while the cytochrome b/f complex included five genes. Additionally, the essential Rubisco gene (*rbcL*) was found in the LSC region. Notably, the NADH-dehydrogenase complex was encoded by 11 genes, predominantly located in the SSC region, highlighting its crucial role in energy metabolism. For self-replication, *S. scaphigerum* exhibited a higher number of genes encoding the small ribosomal subunit (15 genes) compared to the large ribosomal subunit (10 genes). Most PCGs were distributed across the LSC and SSC regions, while the IRs mainly contained RNA genes, supporting both functional diversity and genomic stability. Among the non-photosynthetic genes, *matK* and *cemA* were notable, with the latter encoding the chloroplast envelope membrane protein. The presence of conserved open reading frames further suggested functional roles that remained to be fully elucidated. The gene contents in *S. scaphigerum* was found to be consistent with previous studies[29].

Table 2. List of PCGs present in *Scaphium scaphigerum*

Category	Gene groups	Gene names
Genes for photosynthesis	Photosystem I	*psaA, psaB, psaC, psaI, psaJ*
	Photosystem II	*psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ*
	ATP synthase	*atpA, atpB, atpE, atpF, atpH, atpI*
	Cytochrome b/f complex	*petA, petB, petG, petL, petN*
	Rubisco	*rbcL*
	NADH-dehydrogenase	*ndhA, ndhB*(×2), *ndhC, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK*
Self-replication	Ribosome: small subunit	*rps2, rps3, rps4, rps7*(×2), *rps8, rps11, rps12*(×3), *rps14, rps15, rps16, rps18, rps19*
	Ribosome: large subunit	*rpl2*(×2), *rpl14, rpl20, rpl22, rpl23*(×2), *rpl32, rpl33, rpl36*
	DNA dependent RNA polymerase	*rpoA, rpoB, rpoC1, rpoC2*
Other genes	Maturase	*matK*
	Envelop membrane protein	*cemA*
	Acetyl-CoA-carboxylase	*accD*
	C-type cytochrom synthesis gene	*ccsA*
	Protease	*clpP*
Unknown	Conserved open reading frames	*ycf1, ycf2*(×3), *ycf3, ycf4, ycf15*(×2)

Repeats and codon usage assessments

MISA-Web server analysis identified 115 SSRs in the plastome of *S. scaphigerum*, predominantly composed of mononucleotides(78) and tetranucleotides (15), with fewer dinucleotides (14), trinucleotides (6), and pentanucleotides (2) (Fig. 3A). In comparison, *Firmiana colorata* had 110 SSRs, while *F. simplex* exhibited the highest number at 124, characterized by 85 mononucleotides, 14 dinucleotides, 3 trinucleotides, 12 tetranucleotides, 9 pentanucleotides and 1 hexanucleotide. *Heritiera fomes* contained the most SSRs overall (133), with 93 mononucleotides, 7 dinucleotides, 6 trinucleotides, 14 tetranucleotides, 11 pentanucleotides, and 2 hexanucleotides. Conversely, *Sterculia nobilis* had 108 SSRs, including one hexanucleotide (Fig. 3A). The REPuter server identified a total of 49 longer repeats in *S. scaphigerum* plastome, including 20 forward repeats, 6 reverse repeats, 22 palindromic repeats, and one complementary repeat (Table 3, Fig. 3B). *F. colorata* exhibited a total of 49 longer repeats, with a predominance of forward repeats (37) but lacked any reverse or complementary repeats, highlighting a distinct pattern in repeat composition. *F. simplex* showed a similar total of 49 repeats, with a slightly lower count of forward repeats (19) but a balanced representation of palindromic (21) and reverse (8) types. *H. fomes* showed a more diverse repeat structure, with 15 forward, 13 reverse, and 19 palindromic, and 2 complementary repeats. Finally, *Sterculia nobilis* had 49 longer repeats as well, comprising 13 forward, 9 reverse, 21 palindromic, and 6 complementary repeats (Fig. 3B). This comparative analysis underscores the diversity and complexity of repeat structures across these species, with *S. scaphigerum* showcasing a balanced representation of repeat types, particularly in palindromic repeats, which may have specific evolutionary and functional implications[32].

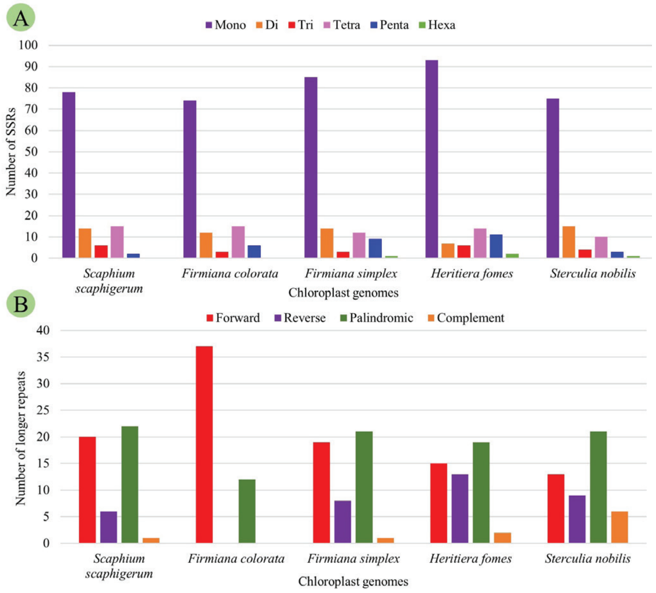


Fig. 3. Repeats present in *S. scaphigerum* and closely related species. A. SSR analysis showing mononucleotides as dominant type in all the taxa. B. Longer repeat analysis showing lack of reverse and complementary repeats in *F. colorata*.

The plastome of *S. scaphigerum* exhibited the usage of 64 unique codons that encode 20 amino acids, with a total codon count of 53,642. Leucine showed the highest Relative Synonymous Codon Usage (RSCU) value of 6.0, indicating a strong bias toward its codons (Fig. 4). Similarly, arginine and serine displayed RSCU values of 6.0, suggesting preferential codon usage for these amino acids. Other amino acids, including valine, proline, threonine, and alanine, had an RSCU of 4.0, indicating balanced codon usage for these residues. Most other amino acids, such as phenylalanine, tyrosine, histidine, glutamine, asparagine, and lysine, exhibited an RSCU values of 2, reflecting moderate codon preference. Stop codons (Ter) displayed an RSCU value of 3, underscoring a potential bias in termination codon selection. This analysis highlights the codon usage bias in *S. scaphigerum*, particularly favoring certain amino acids, which may influence gene expression and translational efficiency. The codon usage pattern observed in *S. scaphigerum* was found to be similar to that of several other species(9,10,33).

Table 3. Evaluation of longer repeats in the plastome of *S. scaphigerum*

Serial No.	Repeat size (bp)	Repeat type	Position 1	Position 2	E-value
1	48	P	79788	79788	9.19E-20
2	35	F	40986	43210	6.17E-12
3	30	P	9850	9881	6.32E-09
4	29	P	8053	48063	2.53E-08
5	27	F	46200	104650	4.04E-07
6	27	P	46200	146262	4.04E-07
7	27	F	105582	105609	4.04E-07
8	27	P	105582	145303	4.04E-07
9	27	P	105609	145330	4.04E-07
10	27	F	145303	145330	4.04E-07
11	26	F	30659	30682	1.62E-06
12	26	R	44219	44219	1.62E-06
13	26	F	51748	84038	1.62E-06
14	26	F	97248	97266	1.62E-06
15	26	P	97248	153647	1.62E-06
16	26	P	97266	153665	1.62E-06
17	26	F	153647	153665	1.62E-06
18	25	R	135239	135239	6.47E-06
19	24	P	10153	10153	2.59E-05
20	24	P	10271	10271	2.59E-05
21	24	P	49893	49893	2.59E-05
22	24	P	132806	132806	2.59E-05
23	23	F	5033	5053	1.04E-04

Table 3. Contd.

Serial No.	Repeat size (bp)	Repeat type	Position 1	Position 2	E-value
24	23	F	31441	131322	1.04E-04
25	23	F	59291	59314	1.04E-04
26	23	F	113786	113818	1.04E-04
27	23	P	113786	137098	1.04E-04
28	23	P	113818	137130	1.04E-04
29	23	F	132296	132318	1.04E-04
30	23	F	137098	137130	1.04E-04
31	22	P	4383	4406	4.14E-04
32	22	F	5462	5482	4.14E-04
33	22	P	32283	32283	4.14E-04
34	22	P	54979	54979	4.14E-04
35	22	P	59343	59369	4.14E-04
36	22	F	61103	61141	4.14E-04
37	22	P	65714	65714	4.14E-04
38	22	F	87848	87869	4.14E-04
39	22	P	116091	116091	4.14E-04
40	22	F	120167	120188	4.14E-04
41	21	C	4792	28028	1.66E-03
42	21	F	8058	37296	1.66E-03
43	21	P	9658	15138	1.66E-03
44	21	P	37296	48066	1.66E-03
45	21	R	44229	44229	1.66E-03
46	21	R	46848	83464	1.66E-03
47	21	R	50039	50039	1.66E-03
48	21	R	55013	55013	1.66E-03
49	21	F	55047	55064	1.66E-03

* P: Palindromic repeat; F: Forward repeat; C: Complementary repeat; R: Reverse repeat.

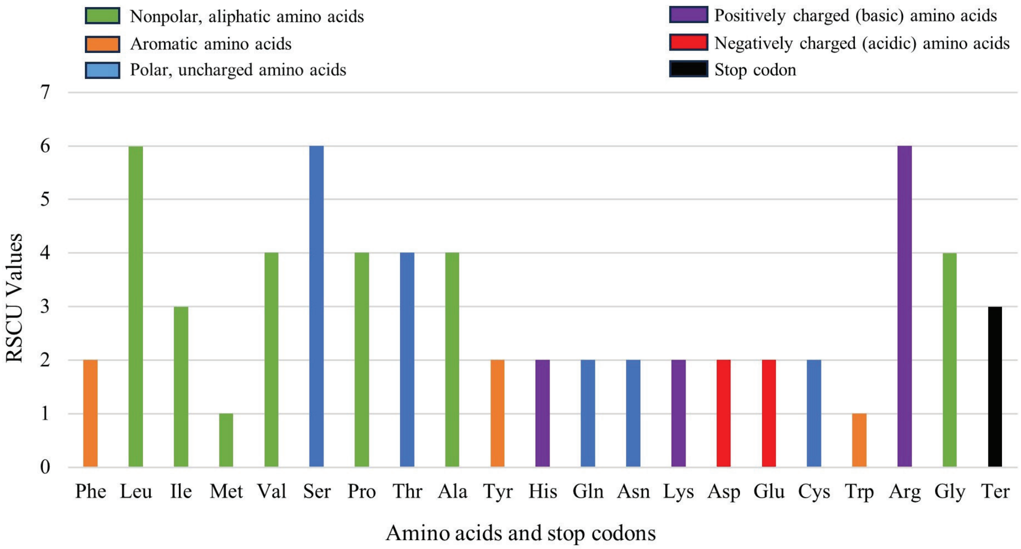


Fig. 4. Codon usage bias in the plastome of *S. scaphigerum.* Amino acids are color-coded according to their diverse classification. Stop codon (Ter) is represented by black color. RSCU value was highest for leucine, serine, and arginine. The lowest RSCU was recorded for methionine and tryptophan.

RNA-editing sites and GC skewness

The RNA-editing analysis identified 66 RNA-editing sites distributed across various genes and regions of the *S. scaphigerum* plastome. The LSC region accounted for the majority of the editing sites, comprising 51% of the total, while the IR region and SSC region contributed 30% and 19%, respectively (Fig. 5). Notably, the *ycf1* gene contained the highest number of editing sites, with 13 sites located in the SSC, indicating significant editing activity in this gene. The *ndhF* gene also showed substantial editing, with 8 sites in the SSC region. In contrast, the LSC region demonstrated multiple genes with single editing site, including *atpB*, *atpF*, and others, while the IR region featured editing in the *ndhB* gene with 7 sites and *ycf2* with 4 sites. The distribution of RNA-editing sites highlights a preference for specific genes in the SSC, whereas the LSC showed a broader range of genes with editing activity, reflecting the complexity of post-transcriptional modifications in the plastome of *S. scaphigerum*.

Skewness analysis revealed that the GC content and GC skew shared a highly similar pattern across all the species examined (Fig. 6). A positive GC skew indicated that guanine (G) was more abundant than cytosine (C) in the analyzed genome regions, while a negative GC skew suggested the opposite. The consistent patterns of GC skew and content across these species suggest a conserved DNA structure and stability, implying that the mechanisms governing guanine and cytosine distribution have been preserved within the Malvaceae family(34).

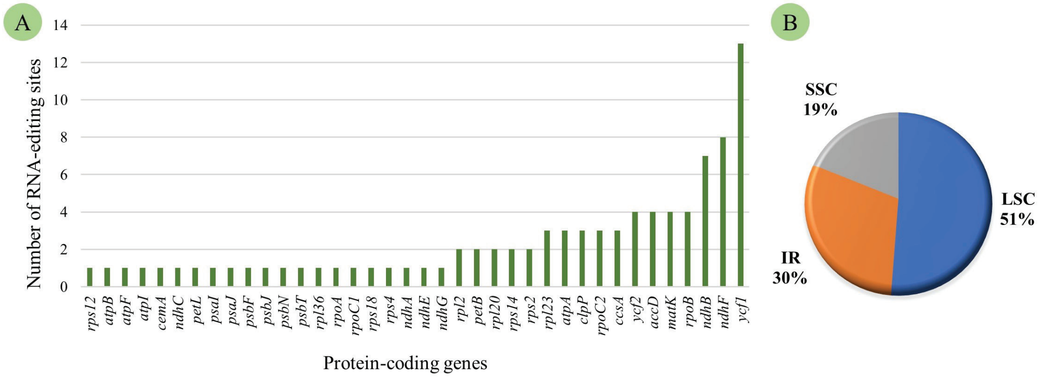


Fig. 5. RNA-editing sites of *S. scaphigerum* plastome. A. Gene-based presentation, B. Compartment-based ratio. Highest number of RNA-editing sites was recorded for the gene *ycf1* (13) that was located in the SSC zone. A total of 21 genes presented single RNA-editing site.

IR contraction and expansion

The junction site analysis of the *S. scaphigerum* plastome revealed a highly similar genomic organization of the quadripartite junctions compared to closely related species (Fig. 7). In *S. scaphigerum*, the LSC region spanned from 88,444 to 91,324 bp, while the SSC region ranged from 10,880 to 20,818 bp, indicating notable variability in the SSC zone across the investigated taxa. This variation highlights significant expansion and contraction events of the IRs. Major IR expansions were observed in *Heritiera littoralis* and *H. fomes*, resulting in the *ndhF* and *ndhA* genes being partially located in both the SSC and IR regions. In contrast, *S. scaphigerum* displayed IR contraction, with the *ndhF* gene entirely located in the SSC region and not extending into the IRs. All studied taxa, except for *H. littoralis* and *H. fomes*, exhibited the *rps19* at the LSC/IRb border, further illustrating the impact of IR boundary shifts on gene location. Notably, the plastome of *S. scaphigerum* demonstrated a closer structural resemblance to *Firmiana* species than to other studied taxa, reflecting its phylogenetic proximity. These findings underscore the critical role of IR expansion and contraction in shaping plastome architecture. Such events influence gene arrangement and plastome stability, offering valuable insights into evolutionary divergence and genome plasticity among taxa. The IR boundaries are particularly important, as they impact the preservation of essential genes, potentially contributing to adaptive evolution and species differentiation within the *Malvaceae* family(10,35,36).

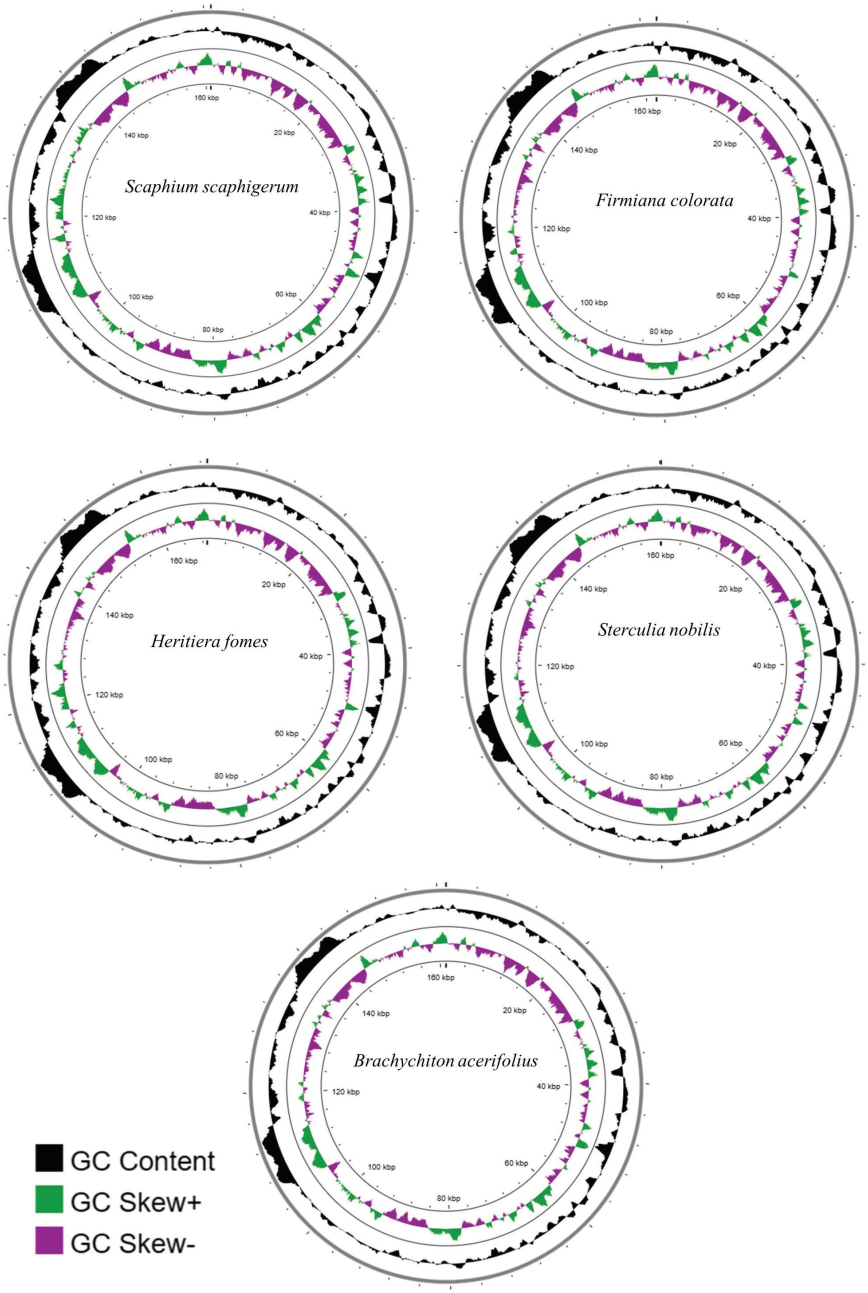


Fig. 6. GC content and skewness analysis of the *S. scaphigerum* plastome along with closely related species. GC content, positive skewness, and negative skewness were indicated by black, green, and violet colors. A high degree of resemblance was observed in the genomic arrangements of the tested taxa *S. scaphigerum*, *Firmiana colorata*, *Heritiera fomes*, *Sterculia nobilis*, and *Brachychiton acerifolius.*

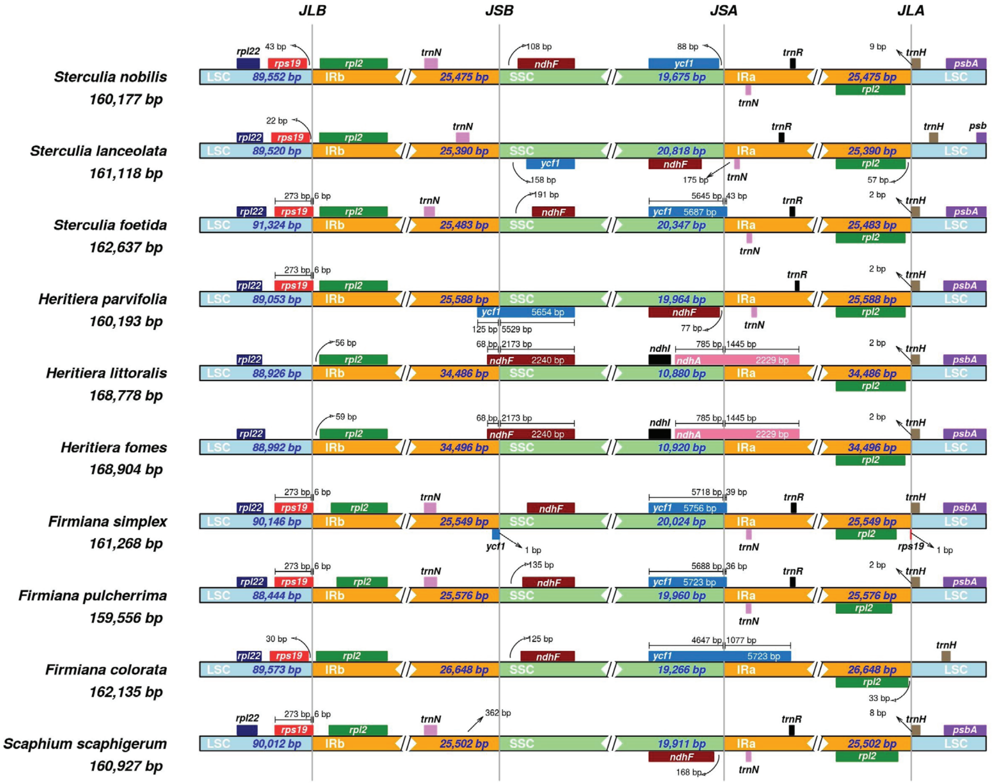


Fig. 7. Expansion and contraction analysis of IRs in the plastome of *S. scaphigerum* and closely related taxa. The numbers placed above or next to the colored genes indicate the distances between each gene and the border edges. The four junction sites have been denoted by JLA (Junction site between LSC and IRa), JSA (Junction site between SSC and IRa), JSB (Junction site between SSC and IRb), and JLB (Junction site between LSC and IRb).

Genome rearrangement and collinearity analysis

The mVISTA analysis revealed that the IRa and IRb regions of the *S. scaphigerum* plastome showed lower levels of genome divergence compared to the LSC and SSC regions. Coding sequences within the genome were more conserved, whereas non-coding regions exhibited higher variability (Fig. 8). Table 4 provides a comparative overview of the plastomes of related taxa analyzed in the present investigation. Plastome-wide alignment revealed locally collinear blocks (LCBs) with significant similarities across the genomes (Fig. 9). The alignment illustrated gene order and arrangements through multi-colored blocks: white for protein-coding genes (PCGs), black for tRNAs, green for intron-containing tRNAs, and red for rRNAs. The high degree of similarity between *S. scaphigerum* and closely related taxa further validated the accuracy of its plastome assembly and annotation. These findings are consistent with previous whole-genome alignment studies[37,38], reinforcing the reliability of the assembly and the comparative analysis.

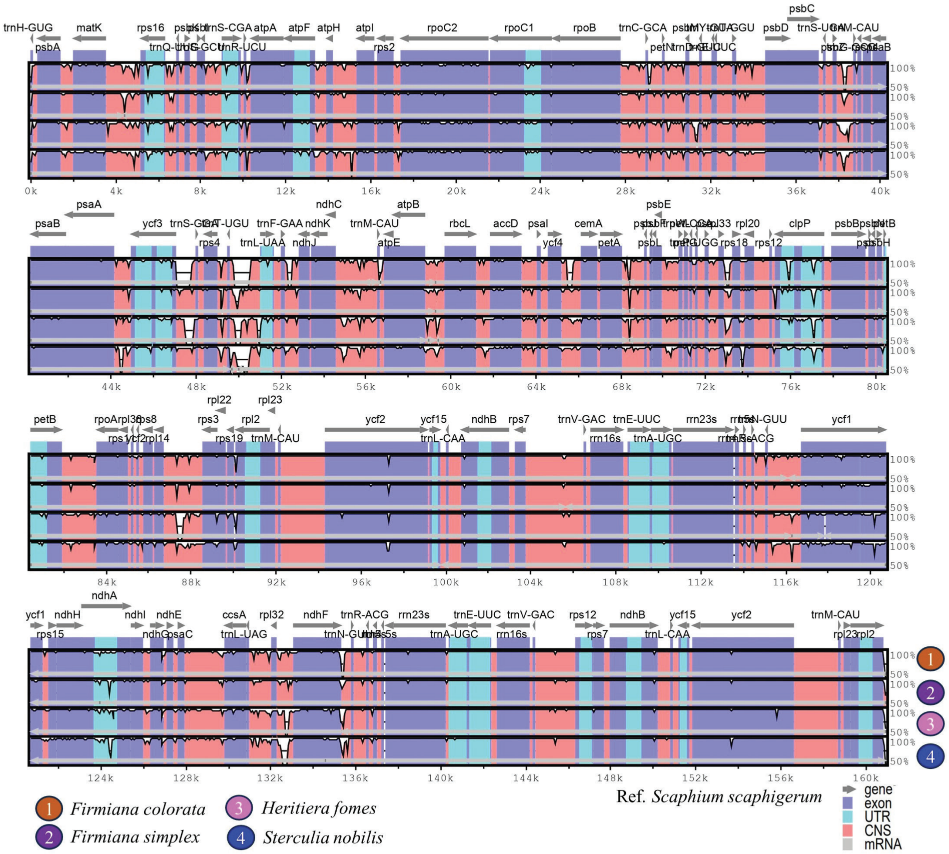


Fig. 8. Genomic variability of *S. scaphigerum* Cp genome in relation with other closely related taxa. Short arrows adjacent to genes indicate the transcription direction of the genes. Exon, untranslated region (UTR), conserved non-coding sequence (CNS), and mRNA have been indicated by various colors.

The collinearity analysis conducted using the Circoletto server revealed a high degree of synteny between *S. scaphigerum* and other closely related taxa (Fig. 10). The absence of significant genomic rearrangements among the studied species underscores their structural similarity and genomic integrity. Among the taxa analyzed, *S. scaphigerum* showed the highest sequence similarity with *Firmiana simplex*, further supporting their close phylogenetic relationship. This high synteny and sequence conservation not only affirm the evolutionary proximity of these taxa but also highlights the robustness and validity of the plastome assembly of *S. scaphigerum*(9). The accurate plastome construction enhances its utility for comparative genomic studies, phylogenetic assessments, and evolutionary investigations, providing a reliable reference for future research on *Scaphium* species and related taxa within the family.

Table 4. Comparative analysis of chloroplast genomes studied in the present investigation

Taxa	Accessions	Tribe/Family	Plastome (bp)	GC (%)	PCGs	tRNAs	rRNAs	Total Genes
Aphanamixis polystachya	NC_048996.1	Meliaceae	160,236	37.58	85	37	8	130
Brachychiton acerifolius	NC_071829.1	Sterculioideae	161,196	36.96	86	37	8	131
Firmiana calcarea	NC_061661.1	Sterculioideae	161,263	36.87	90	37	8	135
Firmiana colorata	NC_054165.1	Sterculioideae	162,135	37.1	85	37	8	130
Firmiana daweishanensis	NC_083141.1	Sterculioideae	161,194	36.88	88	37	8	133
Firmiana hainanensis	NC_065869.1	Sterculioideae	161,031	36.91	88	37	8	133
Firmiana kwangsiensis	MN786867.1	Sterculioideae	160,836	37.04	88	37	8	133
Firmiana pulcherrima	NC_036395.1	Sterculioideae	159,556	37.13	83	36	8	127
Firmiana simplex	NC_041438.1	Sterculioideae	161,268	36.87	88	37	8	133
Heritiera angustata	NC_037784.1	Sterculioideae	168,953	36.8	85	36	8	129
Heritiera fomes	NC_043924.1	Sterculioideae	168,904	36.83	89	37	8	134
Heritiera javanica	NC_057264.1	Sterculioideae	161,419	37.01	85	37	8	130
Heritiera littoralis	MK033518.1	Sterculioideae	168,778	36.82	88	36	8	132
Melia azedarach	NC_050650.1	Meliaceae	160,393	37.37	89	38	8	135
Scaphium scaphigerum	BK067803	Sterculioideae	160,927	36.89	85	35	8	128
Sterculia foetida	NC_088097.1	Sterculioideae	162,637	36.68	86	37	8	131
Sterculia monosperma	NC_053571.1	Sterculioideae	160,178	37.01	86	37	8	131
Sterculia nobilis	NC_063575.1	Sterculioideae	160,177	37.01	87	37	8	132

Nucleotide diversity

The nucleotide diversity analysis identified some hypervariable sites in the Cp genome of *S. scaphigerum* (Fig. 11). The average nucleotide diversity (Pi) value across the complete plastome was recorded at 0.015735. The *ycf1* and *ndhI* genes exhibited Pi values of 0.11333 and 0.10764, respectively, making them as the most hypervariable sites in the SSC region. Variability in the SSC region was higher than that of LSC and IRs. In the LSC region, the *rps18* gene demonstrated the highest Pi value, followed by *rpl33.* Our findings align closely with the nucleotide diversity observed in the plastomes of *Ipomoea* species, where the SSC region exhibited the greater variability compared to the LSC and IRs(39).

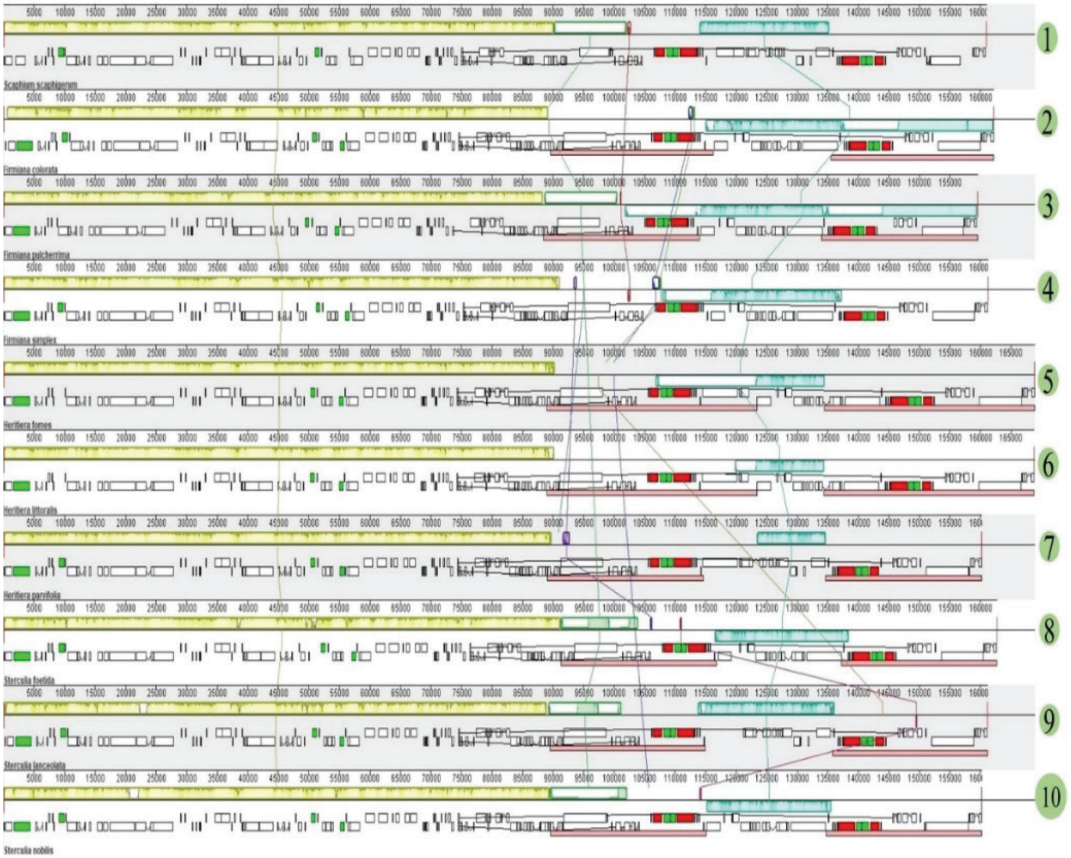


Fig. 9. Whole genome progressive alignment showing genomic similarities of *S. scaphigerum* with other closely related species. 1. *S. scaphigerum*, 2. *Firmiana colorata*, 3. *F. pulcherrima*, 4. *F. simplex*, 5. *Heritiera fomes*, 6. *H. littoralis*, 7. *H. parvifolia*, 8. *Sterculia foetida*, 9. *S. lanceolata*, 10. *S. nobilis*. Protein-coding genes (white) represented almost similar genomic layout along with tRNAs (black), intron-containing tRNAs (green), and rRNAs (red).

Identifying hypervariable genes in the plastome of *S. scaphigerum* is crucial for developing genetic markers or barcodes. The high variability of these genes offers reliable genetic identifiers, facilitating the differentiation of closely related species and subspecies. These hypervariable barcodes enable accurate identification and classification of *Scaphium* species. Furthermore, these markers provide insights into unique mutations within a lineage and aid in elucidating evolutionary relationships, thereby enriching our understanding of the evolutionary dynamics within the family Malvaceae[40].

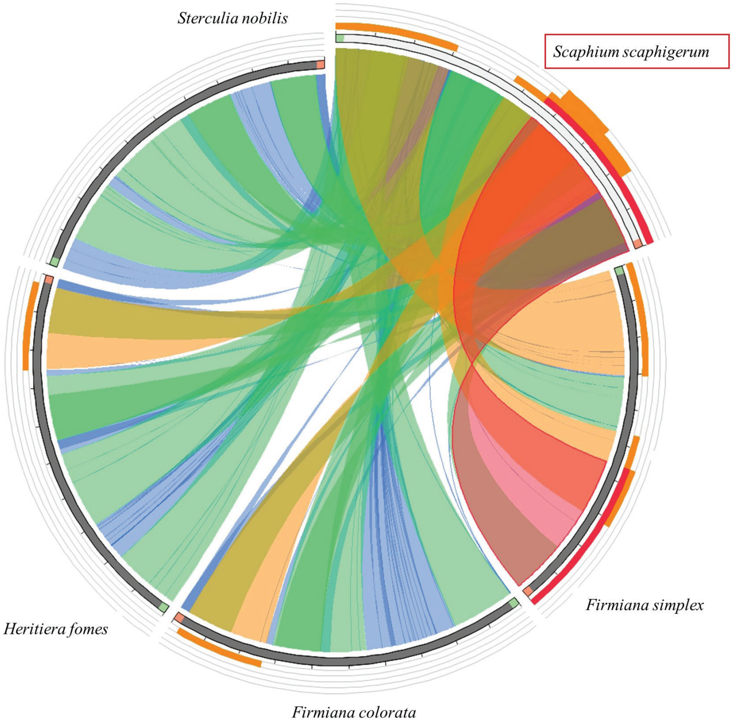


Fig. 10. Synteny analysis of *S. scaphigerum* with other closely related members within Malvaceae. Conserved blocks of genes and genomic regions are highlighted in various colors, showing the degree of structural similarity.

Molecular phylogenetic and dating endeavor

A plastome-wide molecular phylogenetic study revealed the systematic position of *S. scaphigerum* within the subfamily Sterculioideae (Fig. 12). The ML tree demonstrated a well-rooted phylogeny and confirmed the monophyletic origin of Sterculioideae. Bootstrap support was perfect (100%) in almost all branches, underscoring the robust relationships among the taxa. The highest log likelihood recorded for the ML tree was -313525.81. *S. scaphigerum* clustered with the *Firmiana* clade, which aligns well with previously published findings[41]. Wilkie et al.[41] constructed a Maximum-Parsimony (MP) tree using the chloroplast gene *ndhF*, showing *Scaphium* as closely related to the *Firmiana* clade. Our study reinforces that finding by utilizing the entire plastome rather than a single protein-coding gene. All members of *Heritiera* and *Sterculia* were grouped together within their corresponding clades.

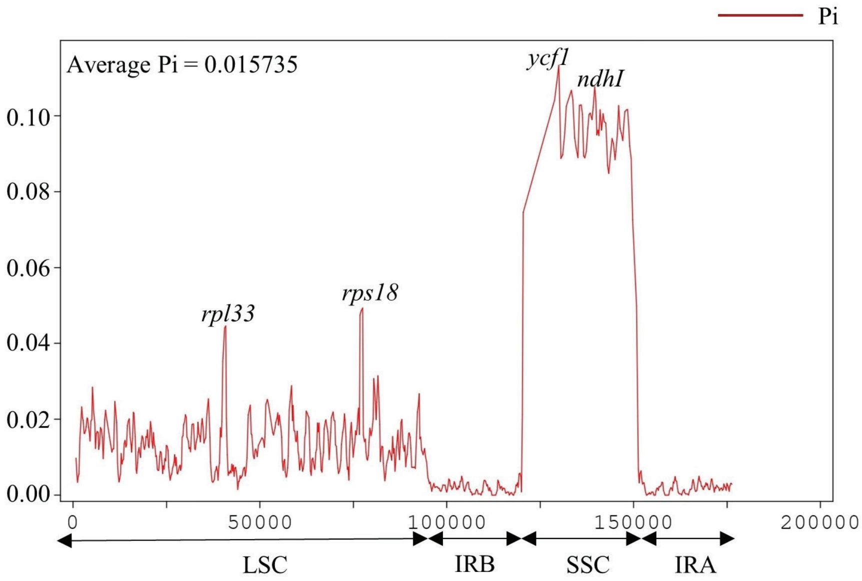


Fig. 11. Nucleotide diversity analysis showing some hypervariable barcodes in the SSC and LSC zones of *S. scaphigerum* plastome. Two most hypervariable sites (*ycf1* and *ndhI*) were located in the SSC zone. Both IR regions represented lower nucleotide diversity. In the LSC zone, *rps18* and *rpl33* genes exhibited higher degree of nucleotide diversity.

For molecular dating analysis, the TimeTree server was executed to identify calibration nodes. Based on available data, the server identified a single calibration node considering *Heritiera* and *Sterculia* genera, with median divergence time of 33 million years ago (MYA), a credibility interval of 30.6 to 65.5 MYA, and an adjusted time estimate of 39 MYA (Fig. 13).

Using this calibration node, a molecular dating tree was constructed that unveiled Sterculioideae to be diverged approximately 51.73 MYA during the Ypresian age. *S. scaphigerum* diverged around 48.23 MYA in the Lutetian age of Cenozoic era (Fig. 14). The genera *Heritiera*, *Sterculia*, and *Firmiana* showed divergence times of 48.14, 38.27, and 30.24 MYA, respectively and originated during the Lutetian, Bartonian, and Rupelian ages of the Cenozoic era correspondingly.

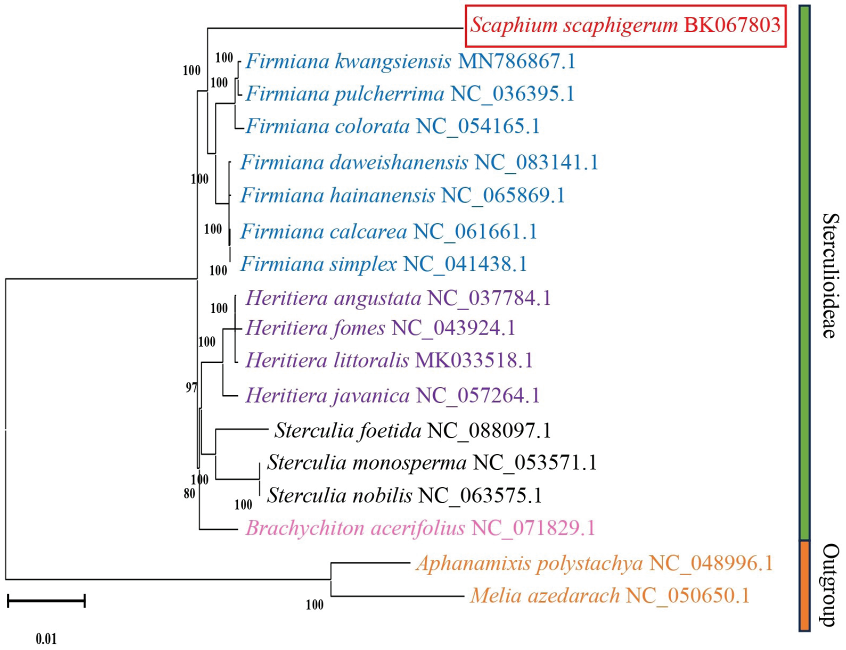
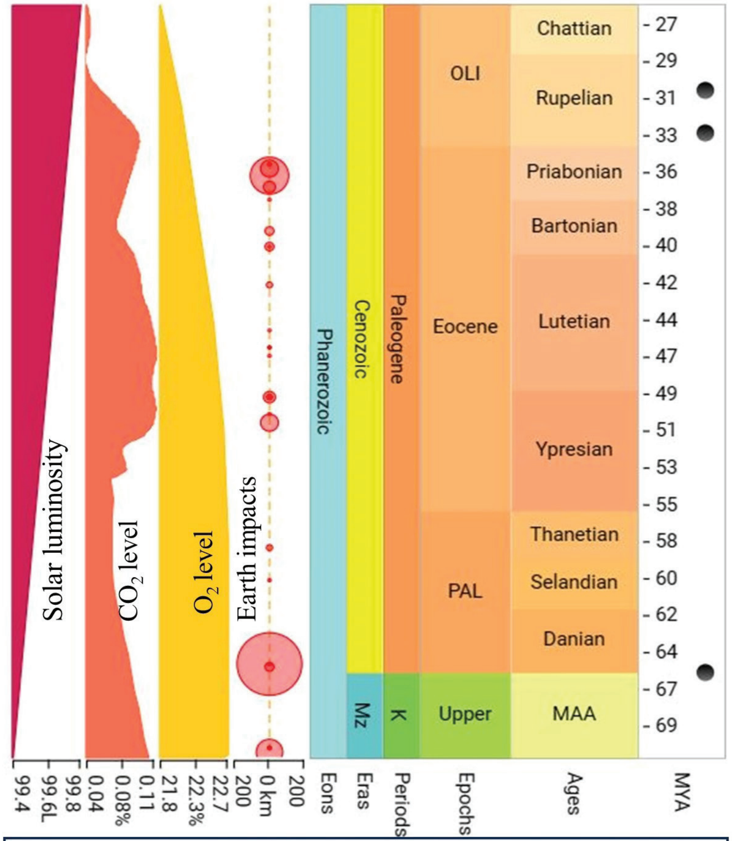

Scaphium scaphigerum BK067803
Firmiana kwangsiensis MN786867.1
Firmiana pulcherrima NC_036395.1
Firmiana colorata NC_054165.1
Firmiana daweishanensis NC_083141.1
Firmiana hainanensis NC_065869.1
Firmiana calcarea NC_061661.1
Firmiana simplex NC_041438.1
Heritiera angustata NC_037784.1
Heritiera fomes NC_043924.1
Heritiera littoralis MK033518.1
Heritiera javanica NC_057264.1
Sterculia foetida NC_088097.1
Sterculia monosperma NC_053571.1
Sterculia nobilis NC_063575.1
Brachychiton acerifolius NC_071829.1
Aphanamixis polystachya NC_048996.1
Melia azedarach NC_050650.1
100
100
100
100
100
100
100
100
100
97
100
80
100
100
0.01
Sterculioideae
Outgroup

Fig. 12. ML tree showing plastome-wide phylogenetic relationships of *S. scaphigerum* within Sterculioideae. The numeric values adjacent to the nodes indicate bootstrap scores. The phylogenetic tree shows a very close affinity between *S. scaphigerum* and members of *Firmiana*. *Aphanamixis polystachya* and *Melia azedarach* were used as outgroup taxa for proper rooting of the tree.

Relying on single gene-based analyses often lead to incomplete phylogenetic trees due to gene-specific evolutionary histories, which can misrepresent relationships, particularly in cases of horizontal gene transfer or lineage-specific adaptations(42). Moreover, molecular dating based on single gene can yield inconsistent divergence estimates, as each gene may have varying rates of evolution influenced by different selective pressures. In contrast, the use of whole plastome data enhances both resolution and accuracy in phylogenetic reconstructions, providing a more reliable temporal framework for identifying key divergence events(43). This knowledge is crucial for understanding the biogeographical distribution and evolutionary trajectories of *S. scaphigerum* and other members of Malvaceae, facilitating the identification of potential genetic markers for conservation and breeding programs.



Heritiera vs. *Sterculia*, Median Time: 33MYA, CI: 30.6 – 65.5 MYA, Adjusted Time: 39 MYA

Fig. 13. Calibration node estimation for the construction of molecular dating tree of *S. scaphigerum.* Eons, eras, periods, epochs, and ages have been represented according to the geological time scale. The smaller black circles indicate molecular time estimates (MYA - million years ago) for the corresponding taxa.

The significance of analyzing publicly available whole genome data for plastome assembly lies in harnessing valuable raw data that might otherwise remain unused in public repositories. The present investigation extracts meaningful biological insights from the publicly available reads, transforming them into a curated Cp genome resource of *S. scaphigerum* for the first time, without requiring additional experimental sequencing. This approach optimizes research efforts and resources, contributing to the broader understanding of plant genetics and conservation. Similar approaches have been employed to construct complete chloroplast genomes utilizing publicly available SRA reads, further supporting our current study(7-9,44).

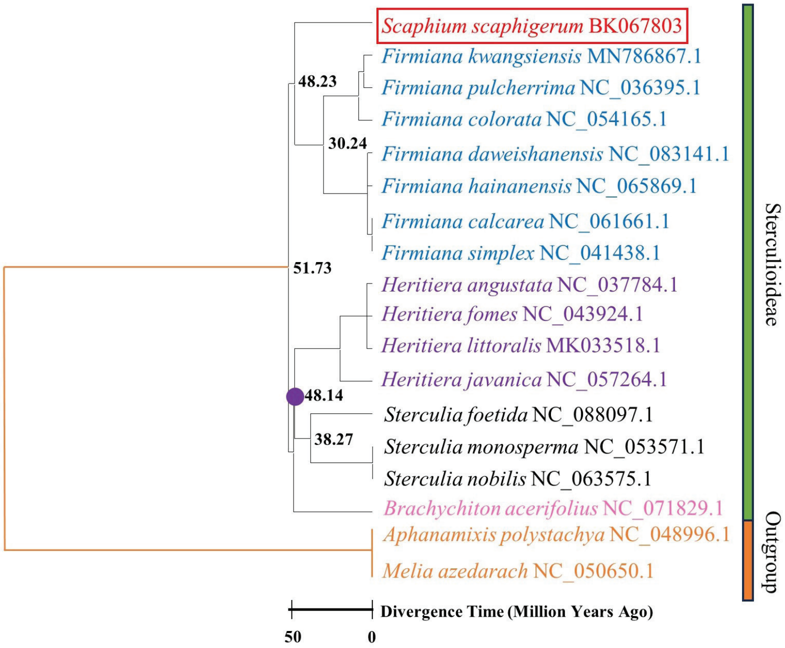

Scaphium scaphigerum BK067803
Firmiana kwangsiensis MN786867.1
Firmiana pulcherrima NC_036395.1
Firmiana colorata NC_054165.1
Firmiana daweishanensis NC_083141.1
Firmiana hainanensis NC_065869.1
Firmiana calcarea NC_061661.1
Firmiana simplex NC_041438.1
Heritiera angustata NC_037784.1
Heritiera fomes NC_043924.1
Heritiera littoralis MK033518.1
Heritiera javanica NC_057264.1
Sterculia foetida NC_088097.1
Sterculia monosperma NC_053571.1
Sterculia nobilis NC_063575.1
Brachychiton acerifolius NC_071829.1
Aphanamixis polystachya NC_048996.1
Melia azedarach NC_050650.1
48.23
30.24
51.73
48.14
38.27
Sterculioideae
Outgroup
Divergence Time (Million Years Ago)
50
0

Fig. 14. Molecular dating tree showing the divergence time according to the geological time scale. The purple circle represents the single calibration node derived from the TimeTree server based on the available taxa. The divergence time was estimated in million years ago (MYA). *S. scaphigerum* originated in 48.23 MYA during the Lutetian age of the Cenozoic era.

The complete chloroplast genome of *S. scaphigerum* (BK067803), representing the first record of this species and currently absent from GenBank, marks a significant advancement in our understanding of the Sterculioideae subfamily. By contributing this genome to the GenBank repository, our findings not only facilitate future research on evolutionary relationships among related taxa but also establish an essential reference for comparative analyses. Additionally, the identification of hypervariable regions within the plastome offers promising avenues for developing DNA barcodes, which can enhance taxonomic resolution and identification in both ecological and medicinal contexts. Therefore, this study will enhance our knowledge of genetic and evolutionary variation within the subfamily Sterculioideae as well as the family Malvaceae.

References

1. Ahmed SS and MO Rahman 2022. Taxonomic revision of the subfamily Sterculioideae Beilschm. in Bangladesh. Bangladesh J. Plant Taxon. **29**(2):373-401.

2. Yamada T, A Itoh, M Kanzaki, T Yamakura, E Suzuki and PS Ashton 2000. Local and geographical distributions for a tropical tree genus, *Scaphium* (Sterculiaceae) in the Far East. Plant Ecol. **148**:23-30.

3. Bennett BC and MJ Balick 2014. Does the name really matter? The importance of botanical nomenclature and plant taxonomy in biomedical research. J. Ethnopharmacol. **152**(3):387-392.

4. Ahmed ZU, MA Hassan, ZNT Begum, M Khondker, SMH Kabir, M Ahmad, ATA Ahmed, AKA Rahman and EU Haque (Eds) 2009. Encyclopedia of Flora and Fauna of Bangladesh, Vol. 10. Angiosperm: Dicotyledons (Ranunculaceae-Zygophyllaceae). Asiatic Society of Bangladesh, Dhaka, pp. 328-351.

5. Phlicharoenphon W, W Gritsanapan, P Peungvicha and P Sithisarn 2017. Determination of antioxidant activity, inhibitory effect to glucose absorption and acute toxicity of *Scaphium scaphigerum* fruit gel powder. J. Health Res. **31**(4):289-296.

6. Daniell H, CS Lin, M Yu and WJ Chang 2016. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. Genome Biol. **17**:1-29.

7. Jongsun P, XI Hong and OH Sang-Hun 2020. Comparative chloroplast genomics and phylogenetic analysis of the *Viburnum dilatatum* complex (Adoxaceae) in Korea. Korean J. Plant Taxon. **50**(1):8-16.

8. Ramadan AM, T Mohammed, KM Al-Ghamdi, AJ Alghamdi and A Atef 2023. The first report describes features of the chloroplast genome of *Withania frutescens*. Saudi J. Biol. Sci. **30**(3):103600.

9. Ahmed SS and MO Rahman 2024. Deciphering the complete chloroplast genome sequence of *Meconopsis torquata* Prain: Insights into genome structure, comparative analysis and phylogenetic relationship. Heliyon **10**(16):e36204.

10. Albediwi AS, MA Ali, MS Alwahibi, SS Ahmed, MO Rahman, SY Kim, MS Elshikh and NM Alsuhaimi 2024. Unveiling the complete chloroplast genome of *Tribulus macropterus* var. *arabicus* (Hosni) Al-Hemaid & J. Thomas: Genome structure, comparative analysis and phylogeny. Bangladesh J. Plant Taxon. **31**(1):1-14.

11. Wei F, D Tang, K Wei, F Qin, L Li, Y Lin, Y Zhu, A Khan, MH Kashif and J Miao 2020. The complete chloroplast genome sequence of the medicinal plant *Sophora tonkinensis*. Sci. Rep. **10**(1):12473.

12. Nguyen HQ, TNL Nguyen, TN Doan, TTN Nguyen, MH Phạm, TL Le, DT Sy, HH Chu and HM Chu 2021. Complete chloroplast genome of novel *Adinandra megaphylla* Hu species: Molecular structure, comparative and phylogenetic analysis. Sci. Rep. **11**(1):11731.

13. Jin JJ, WB Yu, JB Yang, Y Song, CW DePamphilis, TS Yi and DZ Li 2020. GetOrganelle: A fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. Genome Biol. **21**:1-31.

14. Okonechnikov K, O Golosova, M Fursov and Team Ugene 2012. Unipro UGENE: A unified bioinformatics toolkit. Bioinform. **28**(8):1166-1167.

15. Shi L, H Chen, M Jiang, L Wang, X Wu, L Huang and C Liu 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. **47**(W1):W65-W73.

16. Greiner S, P Lehwark and R Bock 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. **47**(W1):W59-W64.

17. Kurtz S, JV Choudhuri, E Ohlebusch, C Schleiermacher, J Stoye and R Giegerich 2001. REPuter: The manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. **29**(22):4633-4642.

18. Beier S, T Thiel, T Münch, U Scholz and M Mascher 2017. MISA-web: A web server for microsatellite prediction. Bioinform. **33**(16):2583-2585.

19. Tamura K, G Stecher and S Kumar 2021. MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Bio. Evol. **38**(7):3022-3027.

20. Lenz H, A. Hein and V Knoop 2018. Plant organelle RNA editing and its specificity factors: Enhancements of analyses and new database features in PREPACT 3.0. BMC Bioinform. **19**:1-18.

21. Grant JR, E Enns, E Marinier, A Mandal, EK Herman, CY Chen, M Graham, GV Domselaar and P Stothard 2023. Proksee: In-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. **51**(W1):W484-W492.

22. Amiryousefi A, J Hyvönen and P Poczai 2018. The chloroplast genome sequence of bittersweet (*Solanum dulcamara*): Plastid genome structure evolution in Solanaceae. PLoS One **13**(4):e0196069.

23. Frazer KA, L Pachter, A Poliakov, EM Rubin and I Dubchak 2004. VISTA: Computational tools for comparative genomics. Nucleic Acids Res. **32**(suppl_2):W273-W279.

24. Darling AC, B Mau, FR Blattner and NT Perna 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. Genome Res. **14**(7):1394-1403.

25. Darzentas N 2010. Circoletto: Visualizing sequence similarity with Circos. Bioinform. **26**(20):p2620.

26. Katoh K, KI Kuma, H Toh and T Miyata 2005. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. **33**(2):511-518.

27. Librado P and J Rozas 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinform. **25**(11):1451-1452.

28. Kumar S, G Stecher, M Suleski and SB Hedges 2017. TimeTree: A resource for timelines, timetrees, and divergence times. Molecular Biol. Evol. **34**(7):1812-1819.

29. Lu Q and W Luo 2022. The complete chloroplast genome of two *Firmiana* species and comparative analysis with other related species. Genetica **150**(6):395-405.

30. Qian J, J Song, H Gao, Y Zhu, J Xu, X Pang, H Yao, C Sun, X Li, C Li, J Liu, H Xu and S Chen 2013. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. PLoS One **8**(2):e57607.

31. Saina JK, AW Gichira, ZZ Li, GW Hu, QF Wang and K Liao 2018. The complete chloroplast genome sequence of *Dodonaea viscosa*: Comparative and phylogenetic analyses. Genetica **146**:101-113.

32. Yi X, L Gao, B Wang, YJ Su and T Wang 2013. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): Evolutionary comparison of *Cephalotaxus* chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms. Genome Biol. Evol. **5**(4):688-698.

33. Samji A, K Eashwarlal, S Shanmugavel, S Kumar and RR Warrier 2023. Chloroplast genome skimming of a potential agroforestry species *Melia dubia* Cav and its comparative phylogenetic analysis with major Meliaceae members. 3 Biotech **13**(1):30.

34. .Wang X, R Zhang, D Wang, C Yang, Y Zhang, M Sui, J Quan, Y Sun, C You and X Shen 2023. Molecular structure and variation characteristics of the plastomes from six *Malus baccata* (L.) Borkh. individuals and comparative genomic analysis with other *Malus* species. Biomolecules **13**(6):962.

35. He S, Y Yang, Z Li, X Wang, Y Guo and H Wu 2020. Comparative analysis of four *Zantedeschia* chloroplast genomes: Expansion and contraction of the IR region, phylogenetic analyses and SSR genetic diversity assessment. PeerJ **8**:e9132.

36. Guo YY, JX Yang, MZ Bai, GQ Zhang and ZJ Liu 2021. The chloroplast genome evolution of Venus slipper (*Paphiopedilum*): IR expansion, SSC contraction, and highly rearranged SSC regions. BMC Plant Biol. **21**(1):248.

37. Henriquez CL, Abdullah, I Ahmed, MM Carlsen, A Zuluaga, TB Croat and MR McKain 2020. Molecular evolution of chloroplast genomes in Monsteroideae (Araceae). Planta **251**:72.

38. Munyao JN, X Dong, JX Yang, EM Mbandi, VO Wanga, MA Oulo, JK Saina, PM Musili and GW Hu 2020. Complete chloroplast genomes of *Chlorophytum comosum* and *Chlorophytum gallabatense*: Genome structures, comparative and phylogenetic analysis. Plants **9**(3):296.

39. Wang Y, J Xu, B Hu, C Dong, J Sun, Z Li, K Ye, F Deng, L Wang, M Aslam, W Lv, Y Qin and Y Cheng 2023. Assembly, annotation, and comparative analysis of *Ipomoea* chloroplast genomes provide insights into the parasitic characteristics of *Cuscuta* species. Front. Plant Sci. **13**:1074697.

40. Breen AL, E Glenn, A Yeager and MS Olson 2009. Nucleotide diversity among natural populations of a North American Poplar (*Populus balsamifera*, Salicaceae). New Phytol. **182**(3):763-773.

41. Wilkie P, A Clark, RT Pennington, M Cheek, C Bayer and CC Wilcock 2006. Phylogenetic relationships within the subfamily Sterculioideae (Malvaceae/Sterculiaceae-Sterculieae) using the chloroplast gene *ndhF*. Syst. Bot. **31**(1):160-170.

42. Ji Y, J Yang, JB Landis, S Wang, Z Yang and Y Zhang 2021. Deciphering the taxonomic delimitation of *Ottelia acuminata* (Hydrocharitaceae) using complete plastomes as super-barcodes. Front. Plant Sci. **12**:681270.

43. Fu CN, CS Wu, LJ Ye, ZQ Mo, J Liu, YW Chang, DZ Li, SM Chaw and LM Gao 2019. Prevalence of isomeric plastomes and effectiveness of plastome super-barcodes in yews (*Taxus*) worldwide. Sci. Rep. **9**(1):2773.

44. Samji A, K Eashwarlal, S Shanmugavel, S Kumar and RR Warrier 2023. Chloroplast genome skimming of a potential agroforestry species *Melia dubia* Cav and its comparative phylogenetic analysis with major Meliaceae members. 3 Biotech. **13**(1):30.

(Manuscript received on 28 August, 2024; accepted on 18 December, 2024)