The complete chloroplast genome of *Hibiscus syriacus* using long-read sequencing: Comparative analysis to examine the evolution of the tribe Hibisceae

Hyunjin Koo[1†], Ah-Young Shin[1,2†], Seongmin Hong[1] and Yong-Min Kim[1,2,3]*

[1]Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, Republic of Korea, [2]Department of Bioinformatics, Korea Research Institute of Bioscience and Biotechnology (KRIBB) School of Bioscience, Korea University of Science and Technology (UST), Daejeon, Republic of Korea, [3]Digital Biotech Innovation Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea

Hibiscus syriacus, a member of the tribe Hibisceae, is considered an important ornamental and medicinal plant in east Asian countries. Here, we sequenced and assembled the complete chloroplast genome of *H. syriacus* var. *Baekdansim* using the PacBio long-read sequencing platform. A quadripartite structure with 161,026 base pairs was obtained, consisting of a pair of inverted repeats (IRA and IRB) with 25,745 base pairs, separated by a large single-copy region of 89,705 base pairs and a short single-copy region of 19,831 base pairs. This chloroplast genome had 79 protein-coding genes, 30 transfer RNA genes, 4 ribosomal RNA genes, and 109 simple sequence repeat regions. Among them, *ndhD* and *rpoC1*, containing traces of RNA-editing events associated with adaptive evolution, were identified by analysis of putative RNA-editing sites. Codon usage analysis revealed a preference for A/U-terminated codons. Furthermore, the codon usage pattern had a clustering tendency similar to that of the phylogenetic analysis of the tribe Hibisceae. This study provides clues for understanding the relationships and refining the taxonomy of the tribe Hibisceae.

KEYWORDS

long-read sequencing platform, complete chloroplast genome assembly, *Hibiscus syriacus*, comparative analysis, Hibisceae

1 Introduction

Hibiscus is one of the most diverse and widespread genera in the Malvaceae tribe Hibisceae (Rizk and Soliman, 2014). The members of the tribe Hibisceae are widely distributed from tropical to temperate regions worldwide (Akpan, 2000). Several species of the tribe Hibisceae are regarded as valuable research crops since they are economically

important for food and medicines and can be utilized as biofuels due to their high biomass content and photosynthetic efficiency (Akpan et al., 2000; Saba et al., 2015; Wu et al., 2017; Cheng et al., 2020). *Hibiscus syriacus*, a member of the tribe Hibisceae, is a flowering shrub that originated in the Korean peninsula and southern China. It is one of the most widely planted ornamental species in temperate zones and is a fast-growing species with attractive white, red, pink, purple, and lavender flowers (Paoletti et al., 2009; Kim et al., 2017). Along with ornamental value, the dried flowers and root bark of *H. syriacus* have been used as a traditional remedy in Oriental countries (Yoo et al., 1998). Particularly, three naphthalene chemicals (syriacusins A–C) and novel pentacyclic triterpene esters identified from the plant's root bark have been used as anthelmintic, antipyretic, and antifungal agents (Yoo et al., 1998; Yun et al., 1999).

Chloroplasts are multifunctional organelles that carry their own genetic sources responsible for photosynthesis, various types of metabolism, and carbon fixation (Li et al., 2019). Chloroplast genomes typically have a quadripartite structure with two copies of inverted repeat (IR) regions separating the large and small single-copy (LSC and SSC, respectively) regions. Most chloroplast genomes range from 120 to 160 kb. Generally, chloroplast genomes of angiosperms contain approximately 120 genes including protein-coding genes, transfer RNA (tRNA), and ribosomal RNA (rRNA) (Daniell et al., 2016). Several mutational events, including mutations, duplications, rearrangements, and gene deletions, occur in chloroplast genomes (Lee et al., 2007). Nevertheless, compared to the nuclear or mitochondrial genome, the chloroplast genome is structurally conserved; hence, it is commonly employed to elucidate the genome evolution and phylogenetic relationships of land plants (Huang et al., 2014; Wu et al., 2017). With the emergence of high-throughput sequencing, the chloroplast genome assemblies of various species of the tribe Hibisceae have been completed (Cheng et al., 2020; Li et al., 2020; Mehmood et al., 2020). Although the phylogenetic relationship among several species of the family Malvaceae was estimated in previous studies, it is insufficient that a comprehensive comparative analysis of the chloroplast genomes in the tribe Hibisceae.

Here, we report the whole chloroplast genome of *H. syriacus* var. *Baekdansim* (hereafter referred to as Baekdansim) using PacBio long-read sequencing data for the first time. Further comparative genome analyses were carried out using the complete chloroplast genomes of other species belonging to the tribe Hibisceae that were obtained from the NCBI database. The findings of this study will be helpful for the development of genetic markers to resolve taxonomic discrepancies or to infer phylogenetic and evolutionary relationships within the tribe Hibisceae.

2 Materials and methods

2.1 Plant material and chloroplast DNA extraction

To isolate high-purity Baekdansim chloroplast DNA from cells, chloroplasts and mitochondria were the first separated from other components, especially nucleus DNA. This step was achieved by homogenizing 5–10 g (fresh weight) of young leaf tissue followed by

a nuclei isolation step according to previous protocols (Zerpa-Catanho et al., 2021). For chloroplast DNA extraction, nuclei removed extract was transferred to 10 mL lysis buffer (50 mM Tris-HCl pH 7.5, 1.4 M NaCl, 20 mM EDTA, pH 8.0, 0.5% SDS) and incubated for 1 h in a 65°C water bath with gentle inversion every 20 min. The supernatant was separated by centrifugation at 3000 rpm for 10 min and transferred to a new tube. RNase A (10 mg/mL) was then added, and the mixture was incubated for 30 min at room temperature. Next, an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the supernatant, and the sample was mixed by gentle inversion for 5 min before centrifugation at 3000 rpm for 10 min. After the aqueous phase was transferred to a new tube, an equal volume of chloroform was added and mixed. The mixture was separated by centrifugation at 3000 rpm for 10 min. The upper, DNA-containing phase was transferred to a new tube, and an equal volume of isopropanol was added to precipitate the DNA, followed by centrifugation at 3000 rpm for 5 min. The DNA pellet was washed with 70% ethanol and resuspended in 100 μL of TE buffer (pH 8.0). Solubilized DNA was stored at 4°C until library preparation.

2.2 Library construction and sequencing

Purified genomic DNA (gDNA) was used for library construction with the SMRTbell Express Template Prep Kit (Pacific Biosciences, Cat. No. 101-357-000). In brief, gDNA was mechanically sheared to an average size of 20 kb using a Covaris g-TUBE device (Part No. 520079). In total, 5 μg of sheared gDNA was damage-repaired and end-repaired using polishing enzymes. Blunt-end adapter ligation was used to create the SMRTbell template. Adapter dimers and contaminants were removed using the AMPure XP bead purification system (Beckman Coulter, Cat. No. A63882). A BluePippin size selection system (Sage Science, Cat. No. BLU0001) was used to size select the SMRTbell template and enrich for fragments > 15 kb. Sequencing primer v4 was annealed to the SMRTbell template, and a DNA polymerase/template complex was created using the Sequel Binding Kit 2.1 (Pacific Biosciences, Cat. No. 101-365-900). An additional AMPure XP purification step was performed to remove excess primer and polymerase prior to sequencing. The library was sequenced on a Sequel instrument using SMRT Cell 1M v2 (Pacific Biosciences), taking one movie of 10 hours per cell with the Sequel Sequencing Kit 2.0 (Pacific Biosciences).

2.3 Genome assembly and annotation

Reads from chloroplasts were extracted by alignment of all reads onto the five chloroplast complete genome assemblies of *Hibiscus* species (*H. syriacus*: NC_026909.1, *H. cannabinus*: NC_045873.1, *H. trionum*: NC_060636.1, *H. rosa-sinensis*: NC_042239.1, and *H. taiwanensis*: NC_045873.1) deposited in the NCBI database (https://www.ncbi.nlm.nih.gov/nucleotide/). Each chloroplast genome was duplicated and concatenated to facilitate the alignment of reads on

the circularized region as suggested by Wang et al. (Wang et al., 2018). Long reads were mapped to chloroplast genomes using minimap2 version 2.24 (Li, 2018). Then, the short reads were mapped using bwa version 0.7.17 (Li and Durbin, 2009). A data set of extracted chloroplast reads was constructed using Unicycler v0.5.0 with the hybrid assembly strategy (Wick et al., 2017). Genome annotation was performed on the GeSeq platform using the complete chloroplast genome (Tillich et al., 2017). Coding sequence (CDS) and rRNAs were predicted by BLAT (Kent, 2002) and HMMER (Finn et al., 2011) search. In addition, the tRNAs were further verified by tRNAscan-SE v2.0.7 (Lowe and Chan, 2016) and ARAGORN v1.2.38 (Laslett and Canback, 2004) with default option. Then, a circular chloroplast map was constructed according to the genome annotation using the online program OGDRAW v1.3.1 (Greiner et al., 2019). The final Baekdansim plastome was deposited in GenBank with accession number OP874596.1. The corresponding circular genome map is shown in Figure 1.

2.4 Repeat sequence identification

Two programs were used to detect repeat motifs. Regarding microsatellites, MISA software (Beier et al., 2017) was used to examine the locations and motifs of simple sequence repeats (SSRs). SSRs were detected using thresholds of 10, 5, 4, 3, 3 and 3 repeat units for mono-, di-, tri-, tetra-, penta-, and hexa-nucleotides, respectively. To identify long repeat motifs, forward, reverse, complementary, and palindromic sequences were determined using REPuter v1.0, with a minimum repetition size of 30 bp and 90% identity (Kurtz et al., 2001).

2.5 Genetic divergence and chloroplast genome comparison

The nucleotide divergence (π) among the 13 species of the tribe Hibisceae was determined using DnaSP v6.0 based on sliding window analysis (Supplementary Table 1) (Rozas et al., 2017). The window length was set to 600 bp with a 100-bp step size. Comprehensive alignments of the complete chloroplast genomes of the tribe Hibisceae were examined using the mVISTA program (Frazer et al., 2004) to reveal interspecific variations. Furthermore, expansion and contraction between the LSC/IRB/SSC/IRA regions at junction sites were identified using IRscope (Amiryousefi et al., 2018). Genes in the chloroplast genomes of 13 species were investigated to determine the presence of introns. Alterations of genes containing intron regions were identified using in-house Python code.

2.6 Codon usage and RNA-editing sites

Relative synonymous codon usage (RSCU) analysis of coding sequences was conducted using MEGA v11.0 (Kumar et al., 1994), and an RSCU value greater than one was regarded as a high codon frequency. The putative RNA-editing sites of the start and stop codons of the coding sequences from species of the tribe Hibisceae were predicted using in-house Python code.

2.7 Phylogenetic analysis

The complete chloroplast genome sequence of Baekdansim, together with those of the other 12 species of the tribe Hibisceae
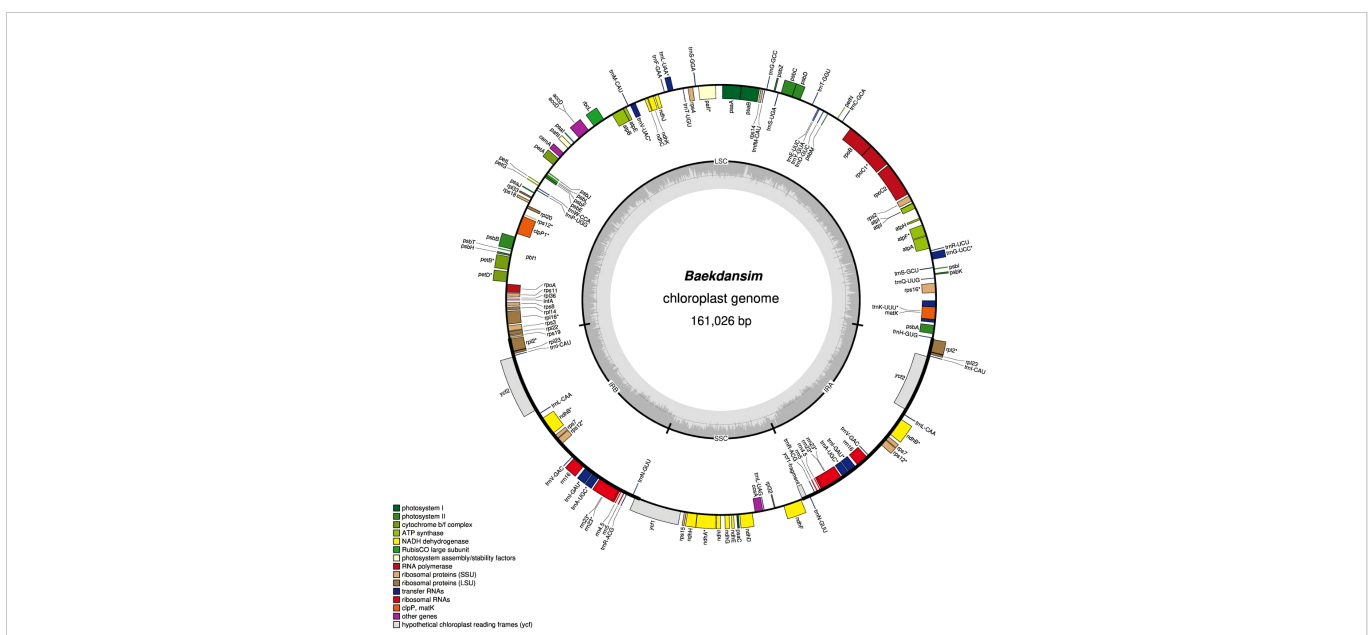

FIGURE 1
Circular chloroplast genome map of *H. syriacus* var. *Baekdansim*. The inner grey circle indicates the proportion of GC in each region. The genes illustrated in the inner circle are transcribed clockwise. Genes corresponding to distinct functional groups are denoted using distinct colors.

available in the NCBI database, were used for comparative and phylogenetic analyses (Supplementary Table 1). The chloroplast sequence of *Gossypium hirsutum* (NC_007944) was also included as an outgroup (Supplementary Table 1). All chloroplast sequences were aligned in MAFFT using default parameters. The best-fit model (K3Pu+F+R4) was estimated using ModelFinder (Kalyaanamoorthy et al., 2017) with Bayesian Information Criterion (BIC) implemented in IQ-TREE v2.0.3 (Minh et al., 2020). Based on the best-fit model (K3Pu+F+R4), we inferred a maximum likelihood tree with 1,000 bootstrap replicates using IQ-TREE. The tree was rooted at midpoint and visualized using FigTree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

3 Results and discussion

3.1 Chloroplast genome assembly

The complete chloroplast genome of Baekdansim was constructed using the PacBio long-read sequencing platform (Figure 1). Due to the high-quality sequence data provided by the PacBio long-read sequencing technique and its capacity to assemble long reads, a single contig and, ultimately, the whole chloroplast genome of *H. syriacus* could be extracted (Chin et al., 2013). The complete chloroplast genome was 161,026 bp and had a quadripartite structure, including a pair of IR regions (IRA and IRB) separated by an LSC (89,705 bp) region and an SSC (19,831 bp) region. The SSC region in the genome assemblies of species of the family Malvaceae in the NCBI database, as determined using short reads, are usually bidirectional. Therefore, the direction of the SSC region was a focus of comparative genome analysis. The genome assembly of Baekdansim consisted of a single contig and was used as a resource to investigate the direction of the SSC region. In a previous study, the primary hypothesis was that the direction of the SSC may have been due to a recombination event between the two IR regions. The alternative hypothesis was that the direction of the SSC region depended on the assembly method; the precise direction of the SSC region was unknown because a short-read-based sequencing platform was used (Cheng et al., 2020). This study showed that the whole chloroplast sequence, which was obtained as a single contig, spanned the whole LSC-IR-SSC area and that the gene order of close species of the family Malvaceae with an inverted SSC structure was exactly reversed. Based on these results, it could be concluded that the SSC direction was changed because of a misassembly induced by constraints of the short read-based sequencing platform. To perform an accurate comparative analysis of species of the tribe Hibisceae, the mis-assembled section was corrected based on the SSC strand derived from long-read sequencing using an in-house Python script.

3.2 Genome structure and gene content

The complete chloroplast genome of Baekdansim contains 113 genes, including 79 protein-coding genes, 4 rRNAs, and 30 tRNAs.

Multiple genes were duplicated in the IR regions, including 5 protein-coding genes (*rpl2, rps7, rpl23, ndhB,* and *ycf2*), 7 tRNAs (*trnA-UGC, trnI-GAU, trnN-GUU, trnV-GAC, trnL-CAA, trnR-ACG,* and *trnI-CAU*), and 4 rRNAs (*rrn5, rrn4.5, rrn23,* and *rrn16*). As generally observed in other angiosperms, 18 intron-containing genes were also detected in the Baekdansim chloroplast genome (Redwan et al., 2015; Mo et al., 2020). Eighteen genes contained one or more introns; of which, 11 encoded proteins (*atpF, clpP1, ndhA, ndhB, pafI, petB, petD, rpoC1, rps16, rpl2,* and *rpl16*), 6 encoded tRNAs *(trnA-UGC, trnG-UCC, trnI-GAU, trnK-UUU, trnL-UAA,* and *trnV-UAC*), and 1 encoded an rRNA (*rrn23*). The *rps12* gene exhibited a trans-spliced form with its 5′ terminal present in the LSC region, and its 3′ end had a single copy present in each of the two IR regions, similar to the patterns observed in other terrestrial plants (Hildebrand et al., 1988; Lee et al., 2019) (Supplementary Table 2). A single copy of *ycf1* was present due to its position in the SSC region instead of in the IR regions. This finding is consistent with a previous study of *Distemonanthus benthamianus*, in which IR contraction was observed (Bai et al., 2021). According to previous reports, the plastome length varies according to the IR length, suggesting that the chloroplast length of *H. syriacus* is also affected by this IR length variation (Zhu et al., 2016; Liang et al., 2022; Lu et al., 2022).

3.3 Repeat analysis

Repeat motifs, which are widely distributed in chloroplast genomes, play an important role in genome evolution (Powell et al., 1995; Yang et al., 2011; Xue et al., 2012; Liu et al., 2018). The number of SSR motifs in the Baekdansim plastome was investigated using MISA software. We identified 109 SSRs (microsatellites) among which 81 (74.3%) consisted exclusively of A/T. Similar to a previous report, the majority of mononucleotide repeats were A/T, and most SSRs consisted of mononucleotide repeats (George et al., 2015). We found 82 (75.2%) mono-, 10 (9.3%) di-, 7 (6.4%) tri-, 6 (5.5%) tetra-, and 4 (3.7%) penta-nucleotides (Supplementary Table 3).

Long-repeat elements are crucial for not only structural variation in chloroplast genomes but also intermolecular recombination, leading to genome diversity (Park et al., 2017; Kong et al., 2021). Complex repeats in Baekdansim were discovered using the REPuter program. The repeat length ranged from 30 bp to 78 bp, which corresponds to the typical range of other plastomes (Greiner et al., 2008; Li et al., 2017). The most abundant repeats were forward repeats, followed by palindromic repeats and reverse repeats (Supplementary Figure 1). These identified repeats and SSRs will be useful for developing molecular markers for genetic diversity and evolution studies.

3.4 Comparison of chloroplast genome structure and nucleotide diversity

Only *Abelmoschus* species contained *rps3b, rps19b,* and *rpl22b,* but other genes were detected in all 13 species of the tribe Hibisceae (Supplementary Figure 2). IRs were more conserved than LSC and SSC sections, while non-coding regions were more divergent than coding regions (Supplementary Figure 3). These results were
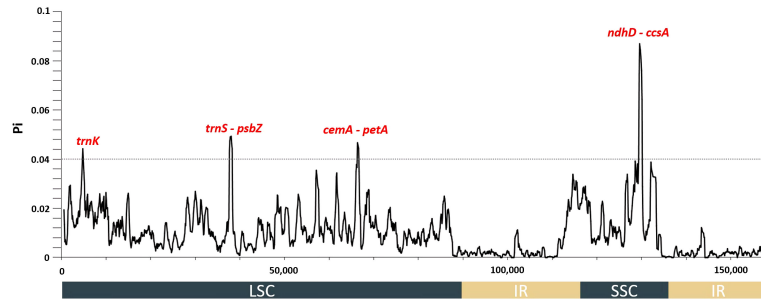
FIGURE 2
Nucleotide diversity values among 13 species of the tribe Hibisceae were calculated using whole plastomes. Mutational hotspots (Pi > 0.04) are denoted above the corresponding gene position.

congruent with findings in other land plant species (Jian et al., 2018; Kim et al., 2020). In addition, the intergenic spacer regions among several gene pairs varied remarkably in chloroplast genomes of the tribe Hibisceae. For instance, these regions differed markedly among *trnH-GUG-psbA*, *trnK-UUU-rps16*, *trnF-GAA-ndhJ*, *atpB-rbcL*, *rps12-trnV-GAC*, *ndhI-ndhG*, and *ndhD-ccsA*. The highest level of nucleotide diversity was identified in a few of these intergenic spacer regions. Collectively, these results suggested these regions might indicate the rapid evolution of the tribe Hibisceae (Figure 2). The nucleotide diversity among the 13 chloroplast genomes of the tribe Hibisceae was calculated. The results indicated four highly divergent hotspots, *trnK, trnS-psbZ, cemA-petA*, and *ndhD-ccsA*, with a threshold of 0.04. All of these hotspots were found in single-copy (LSC and SSC) regions. The most variation was observed in the *ndhD-ccsA* region (0.08703) (Figure 2). It will be important to determine whether these regions

could be employed as DNA barcodes to clarify close relationships within the tribe Hibisceae.

Contraction or expansion of the single-copy and IR regions commonly occurs in various angiosperms (Jian et al., 2018; Li and Zheng, 2018; Henriquez et al., 2020; Guo et al., 2021; Liu et al., 2021). These alterations are considered a major mechanism that causes size variation of the chloroplast genome and evolutionary events (Zhu et al., 2016; Liang et al., 2022; Lu et al., 2022). Four junctions between the two single-copy regions and the two IR regions of 12 representative species of the tribe Hibisceae were thoroughly compared to examine chloroplast genome variation in the tribe Hibisceae (Figure 3; Supplementary Table 4). IR regions are relatively conserved in the *Hibiscus* genus; nonetheless, considerable contraction and expansion occur in the IR/SSC regions. The *ycf1* gene was displaced from the IRB to the SSC region at the IRB/SSC boundary in the chloroplast genomes of *H. syriacus* and *H. rosa-sinensis* by 608 bp and 113 bp, respectively.
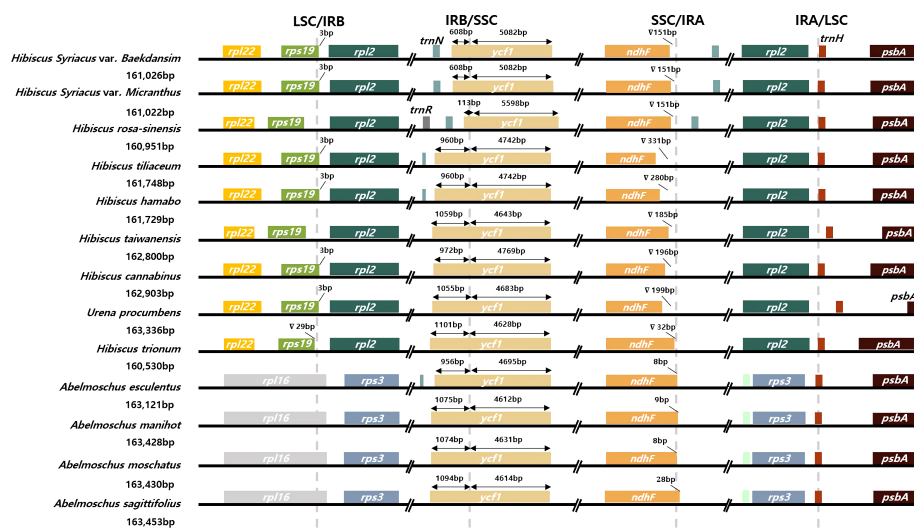

FIGURE 3
Comparison of the junction sites among 13 species of the tribe Hibisceae.

This movement indicates IR contraction in the chloroplast genomes of these species. The *ndhF* gene was shifted from the SSC region to the IRA region, and *rpl16* was shifted from the LSC region to the IRB region in *Abelmoschus* species, according to comparisons between *Hibiscus* and *Abelmoschus* species. The longer chloroplast genome in *Abelmoschus* than in the *Hibiscus* species could be attributed to this IR expansion. In previous reports, shifting of genes to the IR regions or SSC region led to size variation of IR regions in the family Malvaceae. The current study showed that the overall length of the plastome was affected by this size variation shown in previous reports (Dugas et al., 2015; Wang et al., 2016; Guo et al., 2021).

3.5 Putative RNA-editing sites

RNA editing is a post-transcriptional regulation mechanism that can result in the alteration of ribonucleotides at specific sites (Maier et al., 1996). According to previous research, C-to-U conversion is the primary factor responsible for RNA editing (Tsudzuki et al., 2001). Two possible RNA-editing sites were predicted in the start codon of *ndhD* and the stop codon of *rpoC1* in the chloroplast genomes of the tribe Hibisceae (Supplementary Figure 4). In particular, *ndhD* was edited at a high level in *Galium* species. In addition, the start codon (ACG) of *ndhD* in nine species of the family Rubiaceae was affected by an RNA-editing event, which is consistent with the pattern of RNA editing in species of the tribe Hibisceae (Zhang et al., 2019). RNA editing regulates gene expression and has a substantial impact on translation (Maier et al., 1996). RNA editing in the protein-coding region results in codon alterations that lead to amino acid substitution, which may affect the stability of the tertiary structure of proteins (Gommans et al., 2009). Furthermore, these alterations have been related to the generation of genetic diversity, which is a factor in adaptive evolution (Gommans et al., 2009). In tobacco, frequent editing occurs in *ndh*, which encodes the subunits of a plastid NAD(P)H dehydrogenase (Hirose and Sugiura, 1997; Fiebig et al., 2004). Expression of *ndhD* in the tobacco chloroplast, as determined by RNA editing, to create the start codon was greatest in young and photosynthetically active leaves (Hirose et al., 1997). In addition, although *ndh* gene products are dispensable under normal growth conditions, editing is likely essential for the appropriate function of the Ndh protein complex and cyclic electron flow under stress conditions. Fixation of a mutation in a non-essential gene allows plasticity and sufficient time for the evolution of a mutation-compensating editing capacity under moderate selective pressure (Fiebig et al., 2004). Therefore, the occurrence of RNA editing in *ndhD* at the same site in all species of the tribe Hibisceae could be regarded as a result of environmental adaptation. In general, species of the tribe Hibisceae are tolerant of abiotic stresses such as cold, drought, and salt stresses (Zhan et al., 2019; An et al., 2020; Eo et al., 2020; Mahougnon et al., 2021; Chen et al., 2022). Thus, the fixation of RNA editing might have occurred *via* long periods of adaptation to environmental changes during evolution.

3.6 Codon usage pattern and phylogenetic analysis

Many terrestrial plants exhibit codon usage bias, which is considered to play a substantial role in regulating translation dynamics (Du et al., 2020). Recent studies have demonstrated that codon preferences significantly influence the evolution of the chloroplast genome by balancing natural selection and mutational biases (Akashi and Eyre-Walker, 1998; Raubeson et al., 2007; Hershberg and Petrov, 2008). In this study, the RSCU of protein-coding genes in the chloroplast genome of the tribe Hibisceae was investigated and identified (Figure 4). Among the protein-coding codons, the most frequently encoded was leucine, followed by those encoding arginine and alanine; the GAC codon, which encodes aspartic acid, had the lowest usage frequency. If neutral mutations occur at the third codon position, GC and AT would equally present among the codon groups within a chloroplast genome (Zhang et al., 2007). However, most codons showed a bias toward an A/U ending, and these findings are consistent with those observed in other chloroplast genomes (Yan et al., 2019; Du et al., 2020). Previous studies revealed that this unequal usage of nucleotides derived from mutation selection and natural selection was the primary driver of codon bias in angiosperms (Nie et al., 2014; Wang et al., 2020; Zhang et al., 2021). These findings indicate that the high proportion of A/U-ending codons in the chloroplast genome, along with the selective pressure of the chloroplast genome of the tribe Hibisceae, may have driven several degenerate codon biases.

Moreover, species within the tribe Hibisceae were largely clustered into five groups by the RSCU pattern (Group I: *H. syriacus*, Group II: *H. taiwanensis*, *H. tiliaceum*, and *H. hamabo*, Group III: *Urena procumbens* and *H. cannabinus*, Group IV: *H. rosa-sinensis*, Group V: *Abelmoschus sagittifolius*, *A. moschatus*, *A. manihot*, *A. esculentus*, and *H. trionum*) (Figure 4). Phylogenetic analysis was performed on an alignment of the whole chloroplast genome sequences of 13 species of the tribe Hibisceae (Figure 5). It is noteworthy that *H. trionum* formed a clade with *Abelmoschus* species. This tendency was linked to the codon usage pattern of protein-coding genes in the tribe Hibisceae (Figure 4). Despite belonging to a distinct genus, it was assumed that the similarity of codons with other genera may have affected clade formation among other genera. The association between codon usage patterns and the phylogenic topology inferred from the whole chloroplast genome provides strong support for the hypothesis that nucleotide bias induces codon bias.

4 Conclusion

In this study, the complete chloroplast genome of Baekdansim was constructed *via* a long-read sequencing platform for the first time. Through comparisons among species of the tribe Hibisceae, we found that four mutational hot spots could be used to develop DNA barcodes. Furthermore, we identified fixation of candidate RNA-editing sites, a preference for A/U-terminated codons, and a notable codon usage pattern related to phylogenetic relationships.
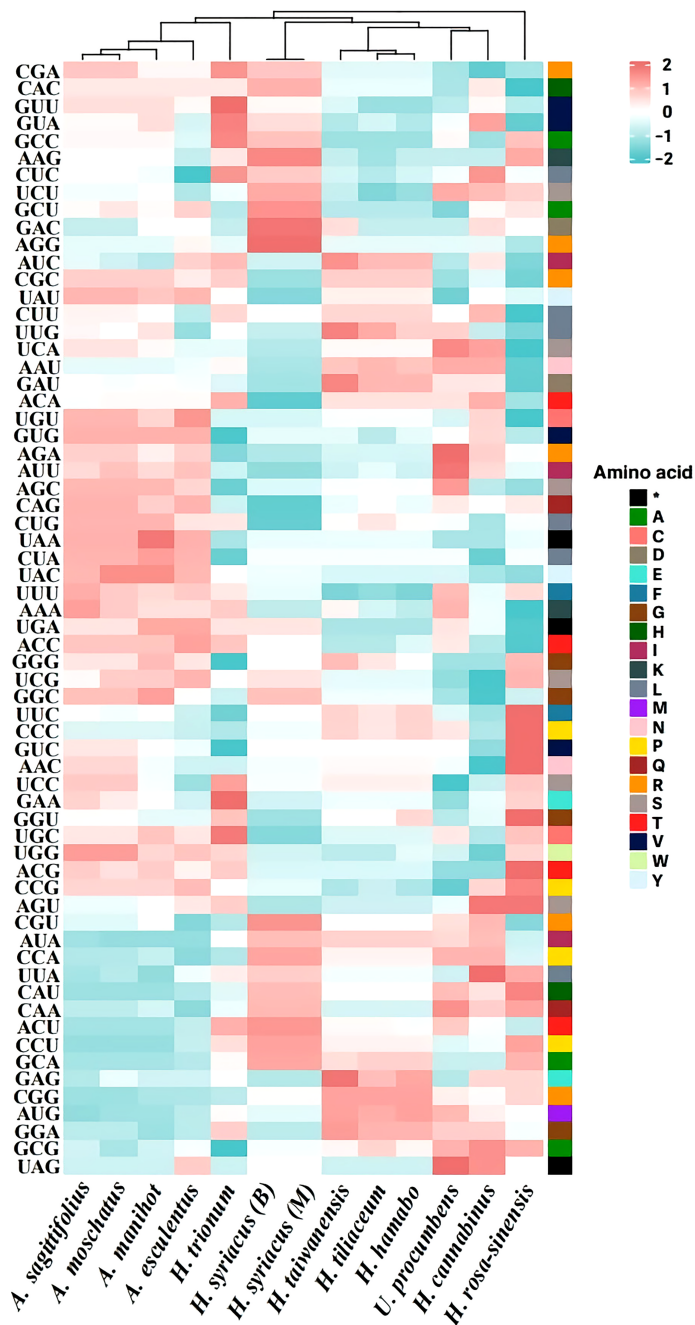
FIGURE 4
Relative synonymous codon usage (RSCU) pattern of chloroplast genes among 13 species of the tribe Hibisceae.
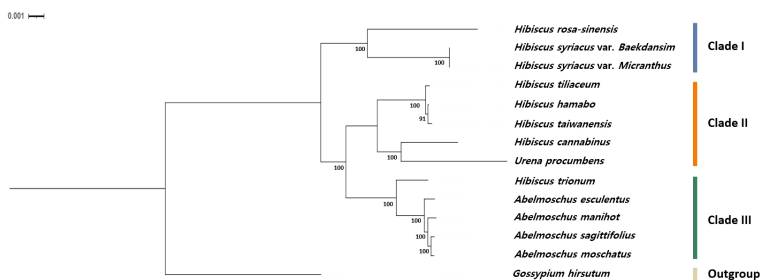

FIGURE 5
Phylogenetic relationships based on the whole chloroplast genomes of 13 species of the tribe Hibisceae. The bootstrap values were based on 1000 replicates and are denoted next to the branches.

Comparison analysis of whole chloroplast genomes of the tribe Hibisceae offers a valuable genomic resource for understanding the evolution and adaptation of this tribe and its relatives.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

HK and Y-MK conceived and designed this study; HK, A-YS, and SH analyzed the data; HK and Y-MK wrote the manuscript; HK, A-YS, and Y-MK revised the manuscript; Y-MK supervised this study. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1111968/full#supplementary-material

References

Akashi, H., and Eyre-Walker, A. (1998). Translational selection and molecular evolution. *Curr. Opin. Genet. Dev.* 8 (6), 688–693. doi: 10.1016/S0959-437X(98)80038-5

Akpan, G. (2000). Cytogenetic characteristics and the breeding system in six *Hibiscus* species. *Theor. Appl. Genet.* 100 (2), 315–318. doi: 10.1007/s001220050041

Amiryousefi, A., Hyvönen, J., and Poczai, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 34 (17), 3030–3031. doi: 10.1093/bioinformatics/bty220

An, X., Jin, G., Luo, X., Chen, C., Li, W., and Zhu, G. (2020). Transcriptome analysis and transcription factors responsive to drought stress in *Hibiscus cannabinus*. *PeerJ* 8, e8470. doi: 10.7717/peerj.8470

Bai, H.-R., Oyebanji, O., Zhang, R., and Yi, T.-S. (2021). Plastid phylogenomic insights into the evolution of subfamily dialioideae (Leguminosae). *Plant Diversity* 43 (1), 27–34. doi: 10.1016/j.pld.2020.06.008

Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33 (16), 2583–2585. doi: 10.1093/bioinformatics/btx198

Cheng, Y., Zhang, L., Qi, J., and Zhang, L. (2020). Complete chloroplast genome sequence of *Hibiscus cannabinus* and comparative analysis of the malvaceae family. *Front. Genet.* 11, 227. doi: 10.3389/fgene.2020.00227

Chen, M., She, Z., Aslam, M., Liu, T., Wang, Z., Qi, J., et al. (2022). Genomic insights of the *WRKY* genes in kenaf (*Hibiscus cannabinus* l.) reveal that *HcWRKY44* improves the plant's tolerance to the salinity stress. *Front. Plant Sci.* 13, 984233. doi: 10.3389/fpls.2022.984233

Chin, C.-S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10 (6), 563–569. doi: 10.1038/nmeth.2474

Daniell, H., Lin, C.-S., Yu, M., and Chang, W.-J. (2016). Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol.* 17 (1), 1–29. doi: 10.1186/s13059-016-1004-2

Dugas, D. V., Hernandez, D., Koenen, E. J., Schwarz, E., Straub, S., Hughes, C. E., et al. (2015). Mimosoid legume plastome evolution: IR expansion, tandem repeat expansions and accelerated rate of evolution in *clpP*. *Sci. Rep.* 5 (1), 1–13. doi: 10.1038/srep16958

Du, X., Zeng, T., Feng, Q., Hu, L., Luo, X., Weng, Q., et al. (2020). The complete chloroplast genome sequence of yellow mustard (*Sinapis alba* l.) and its phylogenetic relationship to other brassicaceae species. *Gene* 731, 144340. doi: 10.1016/j.gene.2020.144340

Eo, H. J., Kwon, H. Y., Da Kim, S., Kang, Y., Park, Y., and Park, G. H. (2020). GC/MS analysis and anti-inflammatory effect of leaf extracts from *Hibiscus syriacus* through inhibition of NF-κB and MAPKs signaling in LPS-stimulated RAW264.7 macrophages. *Plant Biotechnol. Rep.* 14 (5), 539–546. doi: 10.1007/s11816-020-00628-3

Fiebig, A., Stegemann, S., and Bock, R. (2004). Rapid evolution of RNA editing sites in a small non-essential plastid gene. *Nucleic Acids Res.* 32 (12), 3615–3622. doi: 10.1093/nar/gkh695

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39 (suppl_2), W29–W37. doi: 10.1093/nar/gkr367

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32 (suppl_2), W273–W279. doi: 10.1093/nar/gkh458

George, B., Bhatt, B. S., Awasthi, M., George, B., and Singh, A. K. (2015). Comparative analysis of microsatellites in chloroplast genomes of lower and higher plants. *Curr. Genet.* 61 (4), 665–677. doi: 10.1007/s00294-015-0495-9

Gommans, W. M., Mullen, S. P., and Maas, S. (2009). RNA Editing: a driving force for adaptive evolution? *Bioessays* 31 (10), 1137–1145. doi: 10.1002/bies.200900045

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 (W1), W59–W64. doi: 10.1093/nar/gkz238

Greiner, S., Wang, X., Rauwolf, U., Silber, M. V., Mayer, K., Meurer, J., et al. (2008). The complete nucleotide sequences of the five genetically distinct plastid genomes of *Oenothera*, subsection *Oenothera*: I. sequence evaluation and plastome evolution. *Nucleic Acids Res.* 36 (7), 2366–2378. doi: 10.1093/nar/gkn081

Guo, Y.-Y., Yang, J.-X., Bai, M.-Z., Zhang, G.-Q., and Liu, Z.-J. (2021). The chloroplast genome evolution of Venus slipper (*Paphiopedilum*): IR expansion, SSC contraction, and highly rearranged SSC regions. *BMC Plant Biol.* 21 (1), 1–14. doi: 10.1186/s12870-021-03053-y

Henriquez, C. L., Ahmed, I., Carlsen, M. M., Zuluaga, A., Croat, T. B., and McKain, M. R. (2020). Evolutionary dynamics of chloroplast genomes in subfamily aroideae (Araceae). *Genomics* 112 (3), 2349–2360. doi: 10.1016/j.ygeno.2020.01.006

Hershberg, R., and Petrov, D. A. (2008). Selection on codon bias. *Annu. Rev. Genet.* 42, 287–299. doi: 10.1146/annurev.genet.42.110807.091442

Hildebrand, M., Hallick, R. B., Passavant, C. W., and Bourque, D. P. (1988). *Trans*-splicing in chloroplasts: the *rps 12* loci of nicotiana tabacum. *Proc. Natl. Acad. Sci.* 85 (2), 372–376. doi: 10.1073/pnas.85.2.372

Hirose, T., and Sugiura, M. (1997). Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast *ndhD* mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. *EMBO J.* 16 (22), 6804–6811. doi: 10.1093/emboj/16.22.6804

Huang, H., Shi, C., Liu, Y., Mao, S.-Y., and Gao, L.-Z. (2014). Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships. *BMC Evolutionary Biol.* 14 (1), 1–17. doi: 10.1186/1471-2148-14-151

Jian, H.-Y., Zhang, Y.-H., Yan, H.-J., Qiu, X.-Q., Wang, Q.-G., Li, S.-B., et al. (2018). The complete chloroplast genome of a key ancestor of modern roses, *Rosa chinensis* var. *spontanea*, and a comparison with congeneric species. *Molecules* 23 (2), 389. doi: 10.3390/molecules23020389

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14 (6), 587–589. doi: 10.1038/nmeth.4285

Kent, W. J. (2002). BLAT–the BLAST-like alignment tool. *Genome Res.* 12 (4), 656–664. doi: 10.1101/gr.229202

Kim, Y.-M., Kim, S., Koo, N., Shin, A.-Y., Yeom, S.-I., Seo, E., et al. (2017). Genome analysis of *Hibiscus syriacus* provides insights of polyploidization and indeterminate flowering in woody plants. *DNA Res.* 24 (1), 71–80. doi: 10.1093/dnares/dsw049

Kim, Y., Shin, J., Oh, D.-R., Kim, A.-Y., and Choi, C. (2020). Comparative analysis of complete chloroplast genome sequences and insertion-deletion (Indel) polymorphisms to distinguish five *vaccinium* species. *Forests* 11 (9), 927. doi: 10.3390/f11090927

Kong, B. L.-H., Park, H.-S., Lau, T.-W. D., Lin, Z., Yang, T.-J., and Shaw, P.-C. (2021). Comparative analysis and phylogenetic investigation of Hong Kong *Ilex* chloroplast genomes. *Sci. Rep.* 11 (1), 1–13. doi: 10.1038/s41598-021-84705-9

Kumar, S., Tamura, K., and Nei, M. (1994). MEGA: molecular evolutionary genetics analysis software for microcomputers. *Bioinformatics* 10 (2), 189–191. doi: 10.1093/bioinformatics/10.2.189

Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29 (22), 4633–4642. doi: 10.1093/nar/29.22.4633

Laslett, D., and Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32 (1), 11–16. doi: 10.1093/nar/gkh152

Lee, H.-L., Jansen, R. K., Chumley, T. W., and Kim, K.-J. (2007). Gene relocations within chloroplast genomes of *Jasminum* and *Menodora* (Oleaceae) are due to multiple, overlapping inversions. *Mol. Biol. Evol.* 24 (5), 1161–1180. doi: 10.1093/molbev/msm036

Lee, K., Park, S. J., Colas des Francs-Small, C., Whitby, M., Small, I., and Kang, H. (2019). The coordinated action of PPR 4 and EMB 2654 on each intron half mediates *trans*-splicing of *rps12* transcripts in plant chloroplasts. *Plant J.* 100 (6), 1193–1207. doi: 10.1111/tpj.14509

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34 (18), 3094–3100. doi: 10.1093/bioinformatics/bty191

Liang, D., Wang, H., Zhang, J., Zhao, Y., and Wu, F. (2022). Complete chloroplast genome sequence of *Fagus longipetiolata* seemen (Fagaceae): Genome structure, adaptive evolution, and phylogenetic relationships. *Life* 12 (1), 92. doi: 10.3390/life12010092

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, B., Lin, F., Huang, P., Guo, W., and Zheng, Y. (2017). Complete chloroplast genome sequence of *Decaisnea insignis*: Genome organization, genomic resources and comparative analysis. *Sci. Rep.* 7 (1), 1–10. doi: 10.1038/s41598-017-10409-8

Liu, J., Jiang, M., Chen, H., Liu, Y., Liu, C., and Wu, W. (2021). Comparative genome analysis revealed gene inversions, boundary expansions and contractions, and gene loss in the *Stemona sessilifolia* (Miq.) miq. chloroplast genome. *PloS One* 16 (6), e0247736. doi: 10.1371/journal.pone.0247736

Liu, H.-Y., Yu, Y., Deng, Y.-Q., Li, J., Huang, Z.-X., and Zhou, S.-D. (2018). The chloroplast genome of *Lilium henrici*: genome structure and comparative analysis. *Molecules* 23 (6), 1276. doi: 10.3390/molecules23061276

Li, J., Ye, G.-y., Liu, H.-l., and Wang, Z.-h. (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 (11), e0242591. doi: 10.1371/journal.pone.0242591

Li, W., Zhang, C., Guo, X., Liu, Q., and Wang, K. (2019). Complete chloroplast genome of *Camellia japonica* genome structures, comparative and phylogenetic analysis. *PloS One* 14 (5), e0216645. doi: 10.1371/journal.pone.0216645

Li, B., and Zheng, Y. (2018). Dynamic evolution and phylogenomic analysis of the chloroplast genome in schisandraceae. *Sci. Rep.* 8 (1), 1–11. doi: 10.1038/s41598-018-27453-7

Lowe, T. M., and Chan, P. P. (2016). tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 44 (W1), W54–W57. doi: 10.1093/nar/gkw413

Lu, Q.-X., Chang, X., Gao, J., Wu, X., Wu, J., Qi, Z.-C., et al. (2022). Evolutionary comparison of the complete chloroplast genomes in *Convallaria* species and phylogenetic study of asparagaceae. *Genes* 13 (10), 1724. doi: 10.3390/genes13101724

Mahougnon, B. G. G., Julien, K. K., Armel, C. G. M., and Christophe, B. G. (2021). Salinity resistance strategy of okra (*Abelmoschus esculentus* l. moench) cultivars produced in Benin republic. *Int. J. Plant Physiol. Biochem.* 13 (1), 19–29. doi: 10.5897/IJPPB2021.0308

Maier, R. M., Zeitz, P., Kössel, H., Bonnard, G., Gualberto, J. M., and Grienenberger, J. M. (1996). RNA Editing in plant mitochondria and chloroplasts. *Post-transcriptional control Gene Expression Plants* 32 (1-2), 343–365. doi: 10.1007/978-94-009-0353-1_16

Mehmood, F., Shahzadi, I., Waseem, S., Mirza, B., Ahmed, I., and Waheed, M. T. (2020). Chloroplast genome of *Hibiscus rosa-sinensis* (Malvaceae): comparative analyses and identification of mutational hotspots. *Genomics* 112 (1), 581–591. doi: 10.1016/j.ygeno.2019.04.010

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37 (5), 1530–1534. doi: 10.1093/molbev/msaa015

Mo, Z., Lou, W., Chen, Y., Jia, X., Zhai, M., Guo, Z., et al. (2020). The chloroplast genome of *Carya illinoinensis*: genome structure, adaptive evolution, and phylogenetic analysis. *Forests* 11 (2), 207. doi: 10.3390/f11020207

Nie, X., Deng, P., Feng, K., Liu, P., Du, X., You, F. M., et al. (2014). Comparative analysis of codon usage patterns in chloroplast genomes of the asteraceae family. *Plant Mol. Biol. Rep.* 32 (4), 828–840. doi: 10.1007/s11105-013-0691-z

Paoletti, E., Ferrara, A. M., Calatayud, V., Cerveró, J., Giannetti, F., Sanz, M. J., et al. (2009). Deciduous shrubs for ozone bioindication: *Hibiscus syriacus* as an example. *Environ. pollut.* 157 (3), 865–870. doi: 10.1016/j.envpol.2008.11.009

Park, I., Yang, S., Choi, G., Kim, W. J., and Moon, B. C. (2017). The complete chloroplast genome sequences of *Aconitum pseudolaeve* and *Aconitum longecassidatum*, and development of molecular markers for distinguishing species in the *Aconitum* subgenus *Lycoctonum*. *Molecules* 22 (11), 2012. doi: 10.3390/molecules22112012

Powell, W., Morgante, M., McDevitt, R., Vendramin, G., and Rafalski, J. (1995). Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proc. Natl. Acad. Sci.* 92 (17), 7759–7763. doi: 10.1073/pnas.92.17.7759

Raubeson, L. A., Peery, R., Chumley, T. W., Dziubek, C., Fourcade, H. M., Boore, J. L., et al. (2007). Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* 8 (1), 1–27. doi: 10.1186/1471-2164-8-174

Redwan, R., Saidin, A., and Kumar, S. (2015). Complete chloroplast genome sequence of MD-2 pineapple and its comparative analysis among nine other plants from the subclass commelinidae. *BMC Plant Biol.* 15 (1), 1–20. doi: 10.1186/s12870-015-0587-1

Rizk, R. M., and Soliman, M. I. (2014). Biochemical and molecular genetic characterization of some species of family malvaceae, Egypt. *Egyptian J. Basic Appl. Sci.* 1 (3-4), 167–176. doi: 10.1016/j.ejbas.2014.06.002

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34 (12), 3299–3302. doi: 10.1093/molbev/msx248

Saba, N., Jawaid, M., Hakeem, K. R., Paridah, M. T., Khalina, A., and Alothman, O. Y. (2015). Potential of bioenergy production from industrial kenaf (*Hibiscus cannabinus* l.) based on Malaysian perspective. *Renewable Sustain. Energy Rev.* 42, 446–459. doi: 10.1016/j.rser.2014.10.029

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., et al. (2017). GeSeq–versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 45 (W1), W6–W11. doi: 10.1093/nar/gkx391

Tsudzuki, T., Wakasugi, T., and Sugiura, M. (2001). Comparative analysis of RNA editing sites in higher plant chloroplasts. *J. Mol. Evol.* 53 (4), 327–332. doi: 10.1007/s002390010222

Wang, L., Du, H., Wang, D., and Cao, D. (2016). Complete chloroplast genome sequences of *Eucommia ulmoides*: genome structure and evolution. *Tree Genet. Genomes* 12 (1), 1–15. doi: 10.1007/s11295-016-0970-6

Wang, W., Schalamun, M., Morales-Suarez, A., Kainer, D., Schwessinger, B., and Lanfear, R. (2018). Assembly of chloroplast genomes with long-and short-read data: a comparison of approaches using *Eucalyptus pauciflora* as a test case. *BMC Genomics* 19 (1), 1–15. doi: 10.1186/s12864-018-5348-8

Wang, Z., Xu, B., Li, B., Zhou, Q., Wang, G., Jiang, X., et al. (2020). Comparative analysis of codon usage patterns in chloroplast genomes of six *Euphorbiaceae* species. *PeerJ* 8, e8251. doi: 10.7717/peerj.8251

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PloS Comput. Biol.* 13 (6), e1005595. doi: 10.1371/journal.pcbi.1005595

Wu, M., Li, Q., Hu, Z., Li, X., and Chen, S. (2017). The complete amomum kravanh chloroplast genome sequence and phylogenetic analysis of the commelinids. *Molecules* 22 (11), 1875. doi: 10.3390/molecules22111875

Xue, J., Wang, S., and Zhou, S. L. (2012). Polymorphic chloroplast microsatellite loci in *Nelumbo* (Nelumbonaceae). *Am. J. Bot.* 99 (6), e240–e244. doi: 10.3732/ajb.1100547

Yan, C., Du, J., Gao, L., Li, Y., and Hou, X. (2019). The complete chloroplast genome sequence of watercress (*Nasturtium officinale* r. br.): Genome organization, adaptive evolution and phylogenetic relationships in cardamineae. *Gene* 699, 24–36. doi: 10.1016/j.gene.2019.02.075

Yang, A. H., Zhang, J. J., Yao, X. H., and Huang, H. W. (2011). Chloroplast microsatellite markers in *Liriodendron tulipifera* (Magnoliaceae) and cross-species amplification in l. chinense. *Am. J. Bot.* 98 (5), e123–e126. doi: 10.3732/ajb.1000532

Yoo, I.-D., Yun, B.-S., Lee, I.-K., Ryoo, I.-J., Choung, D.-H., and Han, K.-H. (1998). Three naphthalenes from root bark of *Hibiscus syriacus*. *Phytochemistry* 47 (5), 799–802. doi: 10.1016/S0031-9422(97)00674-2

Yun, B.-S., Ryoo, I.-J., Lee, I.-K., Park, K.-H., Choung, D.-H., Han, K.-H., et al. (1999). Two bioactive pentacyclic triterpene esters from the root bark of *Hibiscus syriacus*. *J. Natural products* 62 (5), 764–766. doi: 10.1021/np9804637

Zerpa-Catanho, D., Zhang, X., Song, J., Hernandez, A. G., and Ming, R. (2021). Ultra-long DNA molecule isolation from plant nuclei for ultra-long read genome sequencing. *STAR protocols* 2 (1), 100343.

Zhang, P., Xu, W., Lu, X., and Wang, L. (2021). Analysis of codon usage bias of chloroplast genomes in *Gynostemma* species. *Physiol. Mol. Biol. Plants* 27 (12), 2727–2737. doi: 10.1007/s12298-021-01105-z

Zhang, Y., Zhang, J.-W., Yang, Y., and Li, X.-N. (2019). Structural and comparative analysis of the complete chloroplast genome of a mangrove plant: *Scyphiphora hydrophyllacea* gaertn. f. and related rubiaceae species. *Forests* 10 (11), 1000. doi: 10.3390/f10111000

Zhang, W. J., Zhou, J., Li, Z. F., Wang, L., Gu, X., and Zhong, Y. (2007). Comparative analysis of codon usage patterns among mitochondrion, chloroplast and nuclear genes in *Triticum aestivum* l. *J. Integr. Plant Biol.* 49 (2), 246–254. doi: 10.1111/j.1744-7909.2007.00404.x

Zhan, Y., Wu, Q., Chen, Y., Tang, M., Sun, C., Sun, J., et al. (2019). Comparative proteomic analysis of okra (*Abelmoschus esculentus* l.) seedlings under salt stress. *BMC Genomics* 20 (1), 1–12. doi: 10.1186/s12864-019-5737-7

Zhu, A., Guo, W., Gupta, S., Fan, W., and Mower, J. P. (2016). Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytol.* 209 (4), 1747–1756. doi: 10.1111/nph.13743