BMC Plant Biology

RESEARCH

Open Access

# Assembly and comparative analysis of the complete mitochondrial of *Spodiopogon sagittifolius*, an endemic and protective species from Yunnan, China

Chao Xu[1], Wei Bi[1], Ren-yi Ma[1,2,3], Pin-rong Li[1], Feng Liu[1] and Zhen-wen Liu[1,2,3*]

## Abstract

**Background** *Spodiopogon sagittifolius,* a C4 plant closely related to cultivated crops, is an edible resource and a Class II nationally protected species in China. Endemic to Yunnan, its populations are declining due to habitat destruction, highlighting its resource and conservation importance. Despite its significance, the molecular phylogenetic relationships and genetic mechanisms of adaptive evolution in the genus *Spodiopogon* remain poorly understood.

**Results** We successfully assembled and annotated the first mitochondrial genome of S. *sagittifolius* using HiFi sequencing and the PMAT tool. The genome is 500,699 bp in length with a GC content of 43.15%. Synteny and dN/dS analyses revealed structural and functional conservation of mitochondrial genomes in closely related species, with most protein-coding genes under purifying selection (dN/dS < 1). Notably, nad2 exhibited signs of positive selection (dN/dS = 1.49), indicating potential adaptive evolution. Extensive RNA editing events were detected across 27 protein-coding genes, predominantly involving C-to-U conversions, with synonymous mutations accounting for 49.65% of the edits. Strong codon usage bias favoring A/U-ending codons and the identification of repeat sequences suggest enhanced mitochondrial efficiency and stress adaptation. Phylogenetic analyses confirmed the taxonomic placement of *S. sagittifolius* within the Andropogoneae tribe.

**Conclusions** This study provides the first insights into the mitochondrial genome evolution of *S. sagittifolius*, highlighting key features linked to stress tolerance and adaptive evolution. These findings establish a foundation for its conservation and potential domestication, with implications for crop improvement and ecological resilience.

**Keywords** Andropogoneae, Grasses, C4, Mitogenome, PMAT, RNA editing, Phylogeny analyses, Adaptive evolution, Positive selection

## Background

During the long evolutionary process, the Poaceae family has become one of the most important plant families on Earth, comprising 11,783 species in 12 subfamilies and 789 accepted genera. The subfamily, Panicoideae, includes 3,325 species in 242 genera [1]. *Spodiopogon sagittifolius* Rendle, a C4 plant endemic to Yunnan, China, is listed as a Category II nationally protected plant in China's Directory of National Key Protected Wild Plants (https://www.gov.cn/, accessed July 6, 2024). Despite its

*Correspondence:
Zhen-wen Liu
liuzw2021@163.com
[1] Yunnan Academy of Forestry and Grassland, Kunming 650201, China
[2] Yunnan Key Laboratory of Biodiversity of Gaoligong Mountain, Yunnan Academy of Forestry and Grassland, Kunming 650201, China
[3] Gaoligong Mountain, Forest Ecosystem, Observation and Research Station of Yunnan Province, Kunming 650201, China

ecological importance, this species remains understudied, with only a single chloroplast genome (NC_064989) and molecular fragments available in public databases. To address this gap, we have undertaken the domestication and breeding of wild populations of this species and successfully completed its genome assembly and annotation. The results reveal a genome size of 849.02 Mb, based on a chromosome-level assembly derived from PacBio HiFi reads and Hi-C data (PRJNA1075325, unpublished), confirming it as a diploid species (2n = 2x = 20).

Some wild Poaceae resources could serve as new sources of crops are one of the most compelling and significant in contemporary plant science [2]. Wild germplasm within the Poaceae family represents a valuable genetic reservoir that complements cultivated crop varieties, offering relatively enriched expressions of agronomically significant traits. These include enhanced stress tolerance, robust resistance to pathogenic and herbivorous challenges, superior yield potential, elevated nutritional quality, and optimized photosynthetic efficiency—characteristics that are less prevalent, though not entirely absent, in domesticated germplasm. For instance, *S. formosanus*, a congeneric species of the *S. sagittifolius*, has been domesticated by indigenous Taiwanese people and regarded as a "globally unique" "super future food" due to its higher nutritional and vitamin contents compared to crops like rice, millet, corn, and Job's tears. Moreover, it exhibits drought tolerance, salt tolerance, and resilience to adverse environments. Taiwanese folk traditions have various methods of consuming *S. formosanus*. This wild relative of major crops demonstrates significant value in food, fodder, ecology, and culture. Conserving and exploiting wild Poaceae resources with superior traits is crucial for crop improvement and resource utilization.

Endemic to subtropical mountainous regions, *S. sagittifolius* is critical to Yunnan's ecosystems, where it has been incorporated into ecological restoration initiatives in the dry-hot Jinsha River valley [3]. However, field surveys reveal alarmingly low population sizes and slow regeneration rates. Of the 11 known populations, nine exist outside protected areas, where habitat destruction—caused by human activities, including road construction and quarrying—threatens its survival. Immediate conservation measures are urgently needed. Molecular biology research is essential to provide a scientific foundation for its conservation. Yet, studies on *S. sagittifolius* are still in their infancy. Preliminary surveys and observational studies indicate that this species thrives in a variety of habitats, including subtropical dry-hot valley shrub grasslands, limestone mountain evergreen broadleaf forests, and semi-moist evergreen broadleaf forests, showcasing broad ecological adaptability. In Yunnan Province, *S. sagittifolius* is found in 11 populations, primarily distributed between 101°57′3.20″E and 104°7′33.00″E, and 23°26′56.92″N to 26°45′5.28″N, at altitudes ranging from 1261 to 2286 m, predominantly on sunny slopes. Significant morphological variation exists among populations, with plant heights ranging from 30 to 150 cm.

Plant mitochondrial genomes exhibit remarkable structural and size diversity, spanning an extraordinary range from 16 kb to 11.3 Mb [4], with variations exceeding 100-fold. Within individual organisms, mitochondrial structures manifest diverse morphological configurations, predominantly in linear, circular, and branched forms [5]. Central to cellular function, mitochondria house critical machinery for oxidative phosphorylation essential to ATP synthesis. Their dynamic structural processes, including fission and fusion mechanisms, are pivotal in maintaining cellular homeostasis, managing stress responses, and regulating apoptotic pathways [6, 7]. Intriguingly, while experiencing rapid structural transformations, these genomes simultaneously demonstrate a remarkably slow sequence evolutionary rate, rendering them invaluable phylogenetic markers for assessing evolutionary relationships [8, 9]. Comparative analyses of plant mitochondrial genomes provide critical insights into the balance between structural variability and functional conservation [10], shedding light on genome evolution, adaptive strategies, and energy metabolism. Mitochondrial research has emerged as a cornerstone in comprehending complex biological phenomena, providing critical scientific insights into diverse domains including plant evolutionary history, energy metabolism, adaptive strategies [11], gene flow and hybrid breeding mechanisms [12], cytoplasmic male sterility (CMS) [13], population genetic structures, community adaptations, and responses to environmental changes, particularly climate variations [13].

Currently, mitochondrial genomes for over 30 Poaceae species are available in public databases, yet no mitogenome has been reported for *Spodiopogon*. Here, we assembled and annotated the complete mitochondrial genome of *S. sagittifolius* using third-generation HiFi sequencing. We provide a comprehensive analysis of its genome structure, gene content, codon usage, repeat sequences, RNA editing, phylogenetics, and dN/dS ratios. Our findings offer critical insights into the evolutionary biology and conservation genomics of this species and lay the foundation for its potential horticultural domestication.

## Materials and methods

### Plant material and genome sequencing

Fresh young leaves of *S. sagittifolius* were obtained from the greenhouse of the Yunnan Academy of Forestry and

Grassland Sciences (Kunming, Yunnan Province, China). The original seeds used in this study were collected by Zhen-wen Liu and Chao Xu from Yimen County, Yunnan Province, China (101°57′ E, 24°41′ N). The formal identification of the plant was also conducted by Zhen-wen Liu from the Yunnan Academy of Forestry and Grassland Sciences. A voucher specimen has been deposited in the herbarium of the Yunnan Academy of Forestry and Grassland Sciences under the deposition number DZ1005. The collection of *S. sagittifolius* was authorized by the Yuxi Forestry and Grassland Bureau. The use of this plant in the study complied with all local, national, and international guidelines and legislation regarding research involving plants. High-quality genomic DNA was extracted using the CTAB method [14]. DNA sample quality was checked by agarose gel electrophoresis, and the concentration was measured using Qubit® DNA Assay Kit in Qubit® 3.0 Flurometer (Invitrogen, USA). The qualified samples were sent to Beijing Novogene Co., Ltd. (https://cn.novogene.com/) for the PacBio® HiFi sequencing.

#### Mitochondrial genome assembly and annotation

The whole-genome sequencing PacBio long reads were deposited in the National Center for Biotechnology Information (NCBI) Sequenced Read Archive (SRA) database (https:// www. ncbi. nlm. nih. gov/ sra/; accession numbers SRR27937524). The sample subcommand of the seqkit toolkit was utilized to perform random sampling on a third-generation sequencing FASTQ file, enabling efficient data handling and analysis by extracting a representative subset of the original sequences. The sample subcommand of the seqkit v2.2.0 was utilized to perform random sampling on a third-generation sequencing FASTQ file with the parameters of '-p 0.1 -s 1234'. We successfully assembled the mitochondrial and chloroplast genomes of *S. sagittifolius* using random third-generation reads (~2.11 G) with PMAT v1.5.3 [15, 16] with default settings. PMAT utilized HiFi reads to create the initial assembly graph. Then, long reads (>30 kb) were fragmented into shorter reads (default: 20 kb) using break_long_reads.py and assembled with Newbler software, packaged in runAssembly.sif. Thirdly, seed contigs were selected by aligning all contigs against a local database of 24 conserved mitochondrial genes using BLASTN(v2.6.0) [17]. Candidate seeds were chosen based on length, identity, and coverage. Fourthly, we used a Breadth First Search (BFS) algorithm in seeds_extension.py, seed contigs are extended to include all target mitochondrial contigs, generating an initial assembly graph in GFA format. Next, the assembly graph was cleaned by removing full-path chloroplast and nuclear contigs, tip contigs, and false paths with assembly_graph.py. GFA format files from assembly results of PMAT were visualized using Bandage v0.8.1 [18] to obtain the circular genome manually. Finally, we used whole-genome sequencing short reads from the SRA database (https://www.ncbi.nlm.nih.gov/sra/; accession numbers SRR27937523) to correct the contigs using BWA-MEM [19] and Pilon (v1.23) [20] with default parameters. The coverage depth of the *S. sagittifolius* mitogenome sequences was obtained using BWA-MEM [9] and Samtools (v1.3.1) [21].

For mitogenome annotation, we used the online GeSeq tool [22] and a new custom grogram: IPMGA (http://www.1kmpg.cn/ipmga/; accessed on 27 June 2024). Five reference mitogenome sequences from GenBank: *Saccharum officinarum* (NC_031164.1), *Sorghum bicolor* (NC_008360), *Zea perennis* (NC_008331.1), *Oryza sativa* Indica Group (NC_007886.1) and *Liriodendron tulipifera* (NC_021152.1) were selected for GeSeq annotation, which were used to cross-validate the results from IPMGA. Then, the annotation errors such as start and stop codon or trans-spliced genes of the mitochondrial genome were manually modified using Apollo v1.11.8 [23]. Map of circular *S. sagittifolius* mitogenome was drawn using PMGmap (http://47.96.249.172:16086/drawing/; accessed on 27 June 2024). The chloroplast genome was annotated online using the CPGAVAS2 platform [24]. Finally, the mitogenome and chloroplast genome sequences of *S. sagittifolius* have been deposited in NCBI with accession numbers PP957933 and PP668227, respectively.

#### Analyses of Homologous sequences and RNA editing

To identify the regions shared between mitochondrial and chloroplast genome, mitochondrial and nuclear genome, we compared the three genomes using BLASTN (v0.0.5) [17], and the results were visualized using the Advanced Circos of TBtools [25]. The nuclear genome of *S. sagittifolius* downloaded from NCBI with the accession number of GCA_040448445.1.

To identify RNA editing sites, we performed lncRNA sequencing on *S. sagittifolius*. Total RNA was extracted from leaf tissues and sent to Novogene Co., Ltd. (Beijing, China) for rRNA depletion and strand-specific library preparation, followed by paired-end sequencing (2×150 bp) on the NovaSeq X Plus platform. The mitochondrial genome of *S. sagittifolius* was indexed and used for read mapping with STAR v2.7.1a [26] using default parameters. STAR alignment was performed using standard parameters including BAM output format and sorted by coordinate. RNA editing sites were subsequently identified using REDItools2 [27], and sites with editing coverage≥5 and frequency≥0.1 were retained for further analysis. Through IGV (v2.19.1) visualization [28], we randomly selected 6 genes that contained 98

RNA editing sites. To ensure the accuracy of these sites, PCR primers (Table S11) were designed to flank the editing regions. Genomic DNA (gDNA) and complementary DNA (cDNA), synthesized from leaf RNA using random primers, served as templates for PCR amplification. The resulting PCR products were sequenced via Sanger sequencing. RNA editing events were confirmed by comparing sequence variations between gDNA and cDNA using SnapGene (v6.0.2) [29].

#### Analysis of codon usage and repeat sequences

The protein-coding sequences of the mitogenome were extracted using PhyloSuite v.1.2.3 [30, 31]. MEGAX v10.2.4 [32] was used to perform codon usage bias preference analysis on the PCGs of the mitogenome. Simple sequence repeats (SSRs) in the mitochondrial were identified using the online tool MISA v2.1 [33] (https://webblast.ipk-gatersleben.de/misa/) with parameters setting the SSR motif length from monomeric to hexameric the minimum numbers of repeats as '10, 5, 4, 3, 3, 3' respectively. Additionally, forward (F), reverse (R), palindromic (P), and complementary (C) repeat sequences were identified by REPuter [34] (https://bibiserv.cebitec.uni-bielefeld.de/reputer/) with parameters of '-c -f -p -r -l 30 -h 3 -best 5000'. Tandem repeats (TE) of the mitochondrial genome were identified by the Tandem Repeats Finder [35] (https://tandem.bu.edu/trf/) with parameters of '2, 7, 7, 80, 10, 50, 500'. The boxplots, bar charts, and line graphs in the manuscript were created using ggplot2 [36] within RStudio (2024.04.1 + 748) [37].

#### Phylogenetic trees construction and analyses

In order to identify the phylogenetic position of the 18 Poaceae species representing different main clades in Soreng et al. [1] were selected, and two species from the Cyperaceae were used as outgroups (Table S12). A total of 15 orthologous mitochondrial genes (*atp1, atp9, ccmC, cox1, cox2, cox3, matR, nad1, nad2, nad4, nad6, nad7, nad9, rps12, rps3*) and 26 orthologous chloroplast genes (*atpA, atpE, petA, petD, psbA, psbD, psbI, psbK, psbM, rpl14, rpl16, rpl2, rpl22, rpl33, rpoB, rpoC1, rpoC2, rps11, rps14, rps15, rps19, rps2, rps4, rps7, rps8, ycf4*) among the analyzed species were identified and extracted by PhyloSuite v1.2.3 [30, 31]. Common genes from 21 species were extracted using PhyloSuite. Multiple sequence alignments of these genes were performed with MAFFT v7.471 [38], integrated within PhyloSuite. The aligned common genes were then concatenated. Gblocks v0.91b [39] was employed to optimize the multiple sequence alignment results. Subsequently, phylogenetic tree construction was carried out using IQTREE v2.2.0 [40]. And the resulting trees were visualized, and all parameters were default. We visualized the phylogenetic trees using the EvolView web service (https://www.evolgenius.info/evolview/#/treeview).

#### Comparative analysis of mitogenome in closely related species

To elucidate the structural variations within the mitochondrial genomes of species closely related to *S. sagittifolium*, we selected three species that belong to the same subfamily (Panicoideae) and tribe (Andropogoneae) as *S. sagittifolium* [1]. The selected species included *Chrysopogon zizanioides* (NC_056367.1), *Coix lacryma-jobi* var. *maxima* (MT471099.1), and *Zea luxurians* (NC_008333.1). Pairwise comparisons between the four species were conducted using BLASTN (v0.0.5) [17] with an E-value threshold of $1e^{-10}$ to identify conserved homologous genes. The Multiple Synteny Plot function in TBtools software [25] was then utilized to visualize the genomic collinearity regions. Additionally, dot plots were generated using Gepard-1.40 [41] to compare *S. sagittifolius* with each of the three related species individually.

#### dN/dS ratio analysis

To investigate the adaptive evolution and selective pressures experienced by mitochondrial genomes across different tribes, we selected five species from distinct subfamilies: Oryzoideae, Bambusoideae, Panicoideae, and Chloridoideae [1]. These species were chosen to represent a broad phylogenetic spectrum within the Poaceae family, allowing us to assess the evolutionary dynamics of mitochondrial genomes at different phylogenetic scales. The selected species included *O. coarctata* (MG429050.1) from Oryzoideae, *B. oldhamii* (EU365401.1) from Bambusoideae, *S. sagittifolius*) and *C. zizanioides* (NC_056367.1) from Panicoideae, and *E. indica* (NC_040989.1) from Chloridoideae. We performed dN/dS ratio analysis on these species to evaluate the rates of nonsynonymous (dN) and synonymous (dS) substitutions in their mitochondrial genomes. The dN/dS ratios for protein-coding genes (PCGs) in the mitogenomes were calculated using the Yang & Nielsen method [42] through pairwise comparisons. Instances where dN or dS values were zero were excluded from the statistical analysis.

## Results

#### Assembly and annotation of S. sagittifolius mitochondrial genome

The *S. sagittifolius* is a perennial herb. The whole plant, aerial parts, spikelet, arrow leaf, caryopsis, anther, seed and young plant are shown in Fig. 1A-H.

We obtained the circular mitogenome of *S. sagittifolius*, which is 500,699 bp in length with a GC content of 43.15%. The mitogenome comprises eight unitigs, with

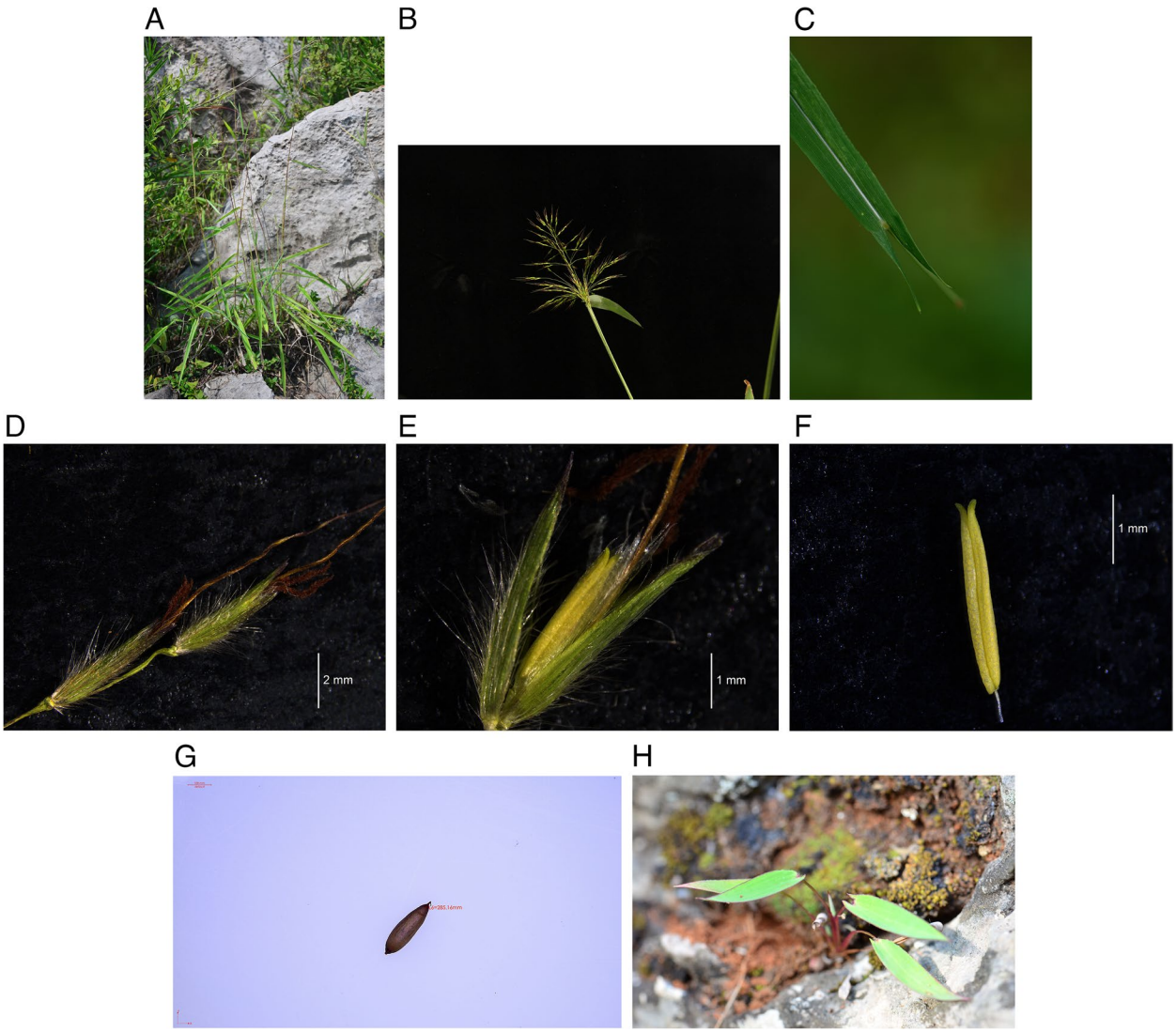


**Fig. 1** Photographs of the aerial parts of *S. sagittifolius*. The images were photographed by Dr. Z.W.Liu from Forestry research institute, Yunnan Academy of Forestry and Grassland Yunnan, China and Dr. L.Q. Jiang from College of Life Science, Henan Normal University, Henan, China, are used in this study with permission. (**A**) The entire plant; (**B**) the spikelet; (**C**) the arrow leaf; (**D**) the spikelet; (**E**) the caryopsis; (**F**) the anther; (**G**) the seed and (**H**) the young plant

their lengths and sequencing depths detailed in Figure S1 and Table S1. The functional categorization and physical locations of the annotated genes are shown in Fig. 2, Table 1 and Table S2. A total of 32 protein-coding genes (PCGs), along with 3 rRNA genes and 24 tRNA genes, were identified. Of these, eight protein-coding genes (PCGs) (*nad1, nad2, nad4, nad5, nad7, ccmFC, cox2,* and *rps3*) contain one or more introns (Table 1). Two genes (*nad1* and *nad5*) require trans-splicing to produce fully assembled, translatable mRNAs, while seven genes (*ccmFC, cox2, nad2, nad4, nad7, rps3,* and *trnF-GAA* ) undergo cis-splicing (Figure S2 A-B). The accuracy of the mitochondrial genome sequence was confirmed by mapping Illumina paired-end short reads, achieving an average coverage of 3,714x, onto the assembled reference mitogenome (Figure S3).

To confirm the loss of genes in Poaceae, we selected 14 species within the Panicoideae subfamily for gene statistical analysis (Table S3). As illustrated in Figure S4, most protein-coding genes (PCGs) are conserved across the mitogenomes of these plants, indicating their crucial roles in mitochondrial function. Notably, *rps14* is present only in the *Oryza sativa Japonica* Group, while *rps19* is retained solely in the *Oryza sativa Japonica* Group, *Alloteropsis semialata*, and *Zea mays* subsp. *m*ays. In the Panicoideae clade, several ribosomal protein large

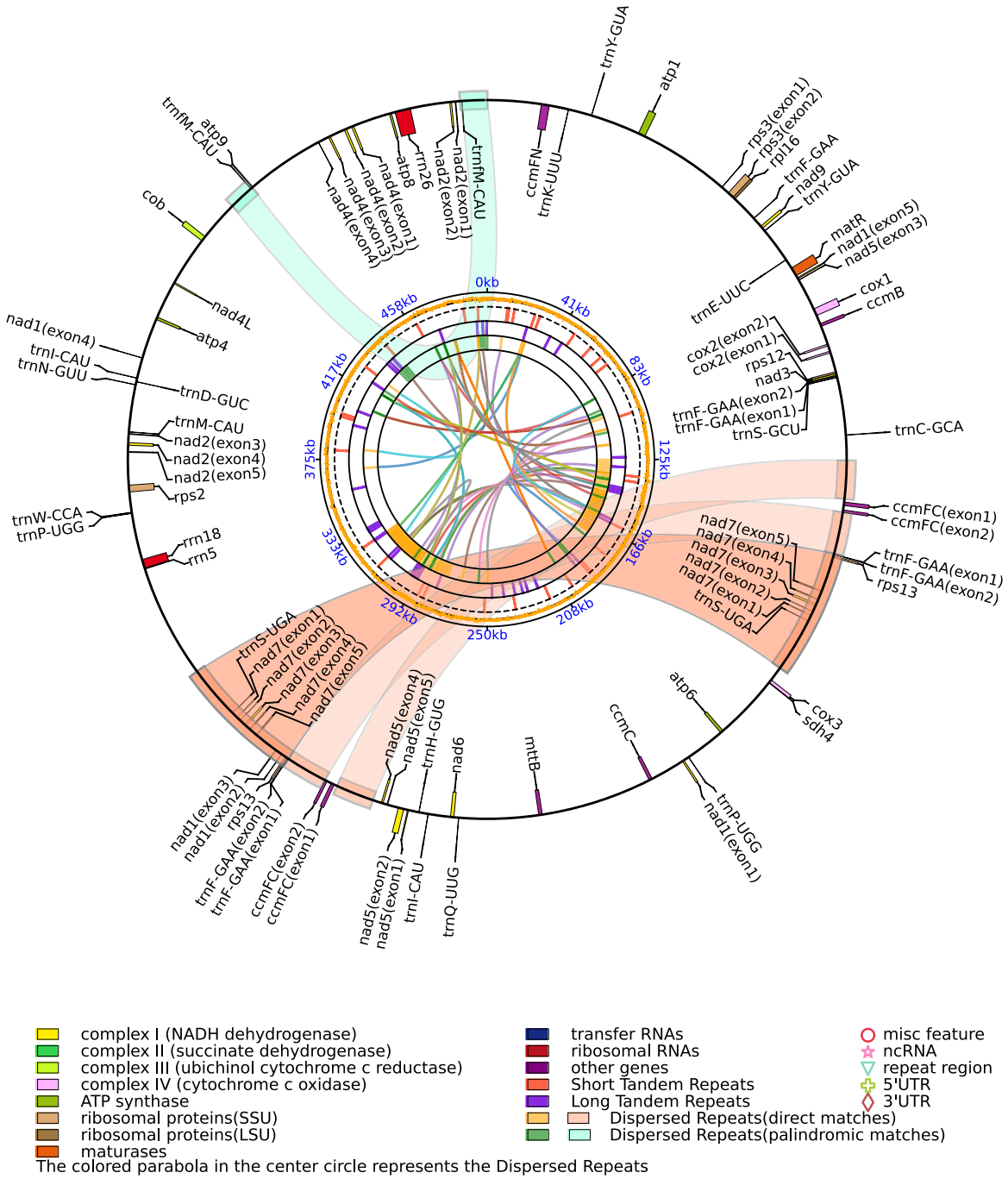


**Fig. 2** Schematic representation of the mitochondrial genome of *S. sagittifolius*. Genes shown outside the circle are on the positive strand, whereas genes shown inside the circle are on the negative strand. The colors represent different functional categories that are described in the legend

**Table 1** Genes predicted in the mitogenomes of *S. sattifolius*

| | Group of Genes | Name of Genes |
|---|---|---|
| Core genes | ATP synthase | atp1, atp4, atp6, atp8, and atp9 |
| | NADH dehydrogenase | nad1[a], nad2[a], nad3, nad4[a] nad4L, nad5[a] nad6, nad7[a](×2) and nad9 |
| | Cytochrome c biogenesis | Cob |
| | Ubiquinol cytochrome c reductase | ccmB, ccmC, ccmFC[a], and ccmFN |
| | Cytochrome c oxidase | cox1, cox2[a], and cox3 |
| | Maturases | matR |
| | Transport membrane protein | mttB |
| Variable genes | Large subunit of ribosome | rpl16 |
| | Small subunit of ribosome | rps2, rps3[a], rps12, and rps13(×2) |
| | Succinate dehydrogenase | sdh4 |
| rRNA genes | Ribosome RNA | rrn5, rrn18 and rrn26 |
| tRNA genes | Transfer RNA | trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA(×4), trnH-GUG, trnI-CAU(×2), trnK-UUU, trnM-CAU, trnfM-CAU(×2), trnN-GUU, trnP-UGG(×2), trnQ-UUG, trnS-GCU, trnS-UGA(×2), trnW-CCA, and trnY-GUA(×2) |

The numbers after the gene names indicate the duplication number.

[a] indicates the genes containing introns

subunit genes (*rpl2, rpl23 and rpl5*) have been widely lost, except for *rpl16*. Additionally, the *sdh4* gene was annotated only in *S. sagittifolius* and *A. semialata*.

#### Repeat sequence analysis

In plant mitochondrial genomes, repetitive sequences, including SSRs (simple sequence repeats), tandem repeats, and TEs (transposable elements), play crucial roles in the analysis of genetic diversity, the characterization of population structure, and the understanding of evolutionary dynamics [43–45]. In the mitogenome of *S. sagittifolius*, 124 simple sequence repeats (SSRs) were identified. Monomeric, dimeric, and trimeric SSRs accounted for 22%, 15%, and 17% of the total, respectively. Tetrameric SSRs were the most common, constituting 40% of the total, while pentameric SSRs were less frequent, comprising only 5% (Fig. 5A-B, Table S4). The mitogenome also contained 405 dispersed repeats (≥30 bp), of which 178 (43.95%) were palindromic and 227 (56.05%) were forward repeats. Most of these duplicated sequences, accounting for 94.07%, were shorter than 100 bp. However, two significant exceptions exceed 5,000 bp: R1, which is 49,049 bp, and R2, which is 6,035 bp (Fig. 3A, Table S5). Additionally, 56 tandem repeats were identified, with details provided in Table S6. These repeats exhibited a wide variability, with counts ranging from 5 to 50.

By conducting a detailed analysis and comparison of the SSR sequences among *S. sagittifolius* and its closely related species within the Andropogoneae tribe [1], we have identified significant variations in the number of SSRs, which range from 114 to 249. Notably, no significant positive correlation was observed between the total SSR count and mitochondrial genome size (Fig. 3B). We also identified five distinct SSRs that differ from those in closely related species: (CG)5(TA)5, (TCA)4, (CGGT)3, (CTAA)3, and (TTCAA)3. The identified SSRs are associated with five intergenic regions: *rpl16—trnF-GAA, sdh4—trnP-UGG, sdh4—atp6, ccmC—mttB, and trnS-UGA—rrn5*. The discovery of these SSR loci within the *Spodiopogon* species indicates that they could serve as valuable genetic markers.

#### Codon usage and dN/dS analysis of PCGs

The codon usage analysis of the 32 protein-coding genes (PCGs) in the mitogenome of *S. sagittifolius* revealed that all 21 amino acids are encoded by a total of 64 distinct codons. This investigation, illustrated in Fig. 4, underscores a pronounced bias towards specific codons in mitochondrial PCGs. Relative synonymous codon usage (RSCU) values greater than 1 identified 28 preferred codons, highlighting their selective usage. Among these, the codon encoding alanine (Ala) exhibited the highest RSCU value of 1.61, indicating a strong preferential bias. Strikingly, among the 28 codons with RSCU values exceeding 1, 92.86% (26 codons) were found to terminate with either A or U, reflecting a notable bias towards A/U-ending codons at the third position. To assess the selection pressure on protein-coding genes (PCGs) in three closely related species of *S. sagittifolius*, we analyzed their dN/dS ratios. As shown in Fig. 5, 23 mitochondrial PCGs, including atp1 (dN/dS = 0.043) (Table S7), exhibited dN/

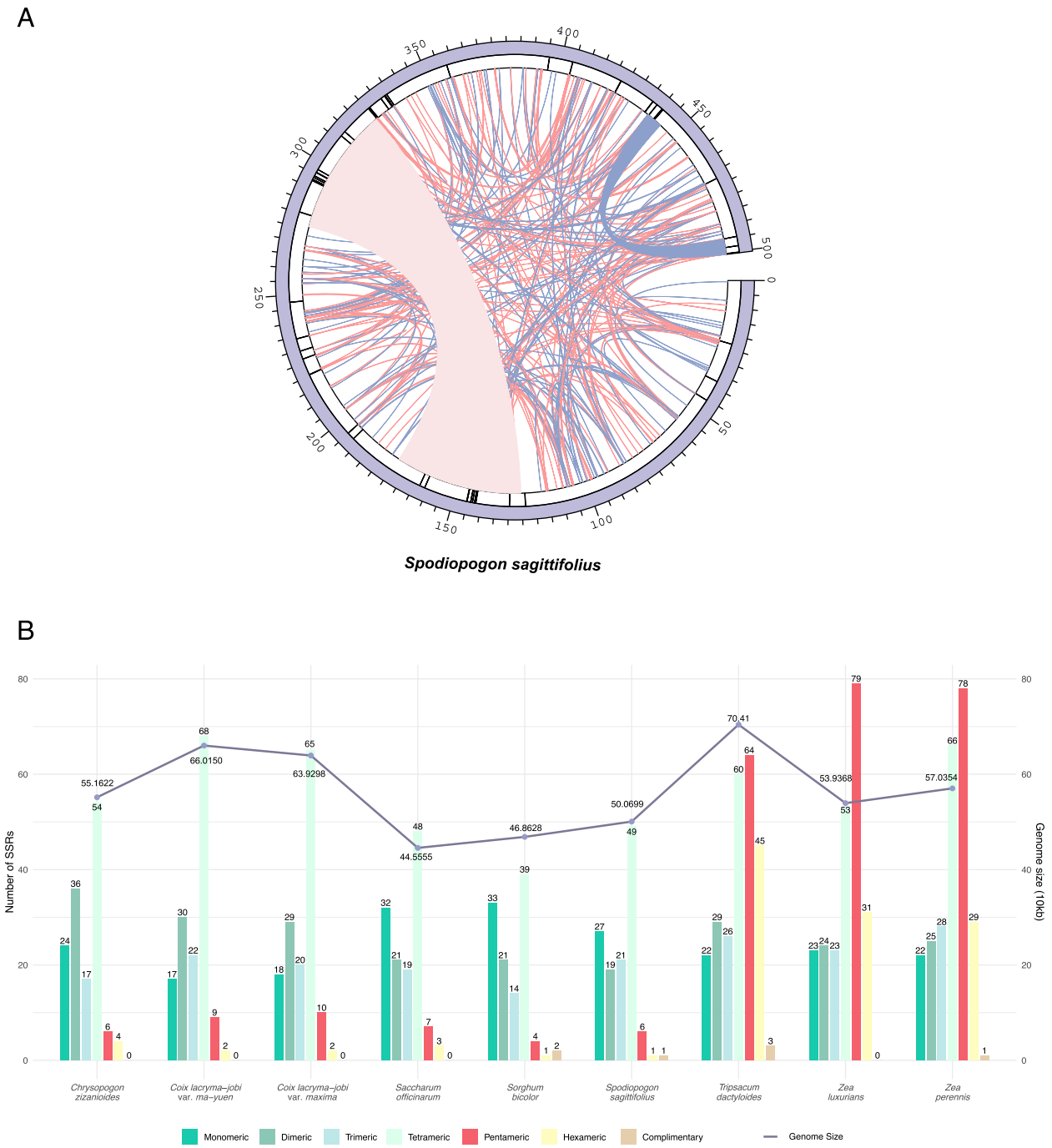


**Fig. 3** Repetitive regions of the mitochondrial genome. (**A**) The outer circle is the tandem repeats, and the inner concatenation is the dispersed repeats among the mitogenome of *S. sagittifolius*. Arcs connect similar repeats within the mitogenome; pink arcs represent 227 forward repeats and the blue represents 178 palindromic repeats (see Figure S4 C, Table S6). Two long repeats are labelled as R1 and R2. (**B**) Comparison of simple sequence repeats (SSRs) and mitochondrial genome sizes in nine closely related species. The x-axis represents different species. Columns in distinct colors correspond to repeat sequences with different types of repeat units. The number of a specific type of repeat is indicated above the respective columns. The size of the mitochondrial genome is connected by a line

dS ratios below 1. In contrast, *nad2* displayed a dN/dS ratio of 1.49, indicating positive selection. The average dN/dS ratios for all PCGs except *nad2*, were below 1, confirming that most genes were under purifying selection and remained highly functional conserved during evolution [6, 46, 47].

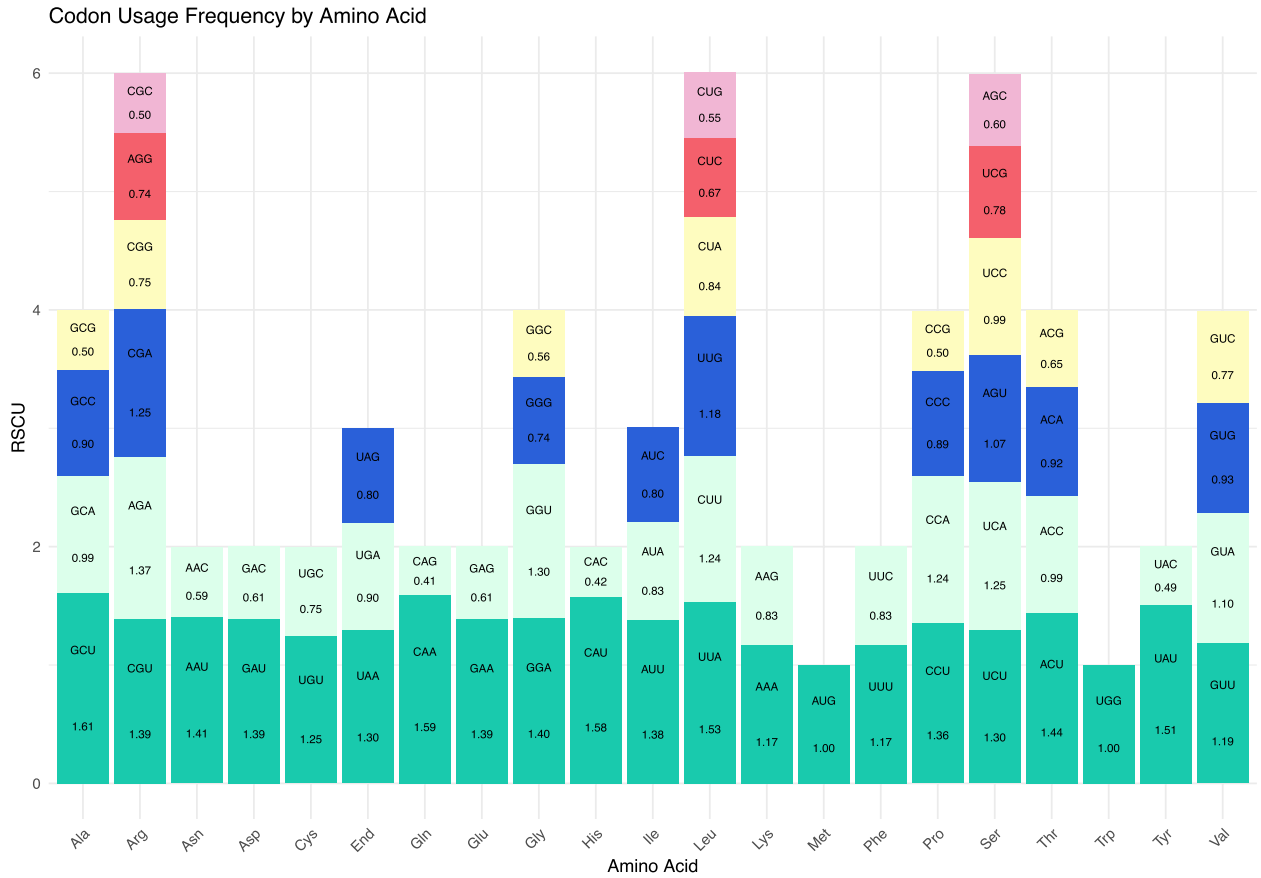


**Fig. 4** The codon content for the 20 amino acids and stop codons in all protein-coding genes

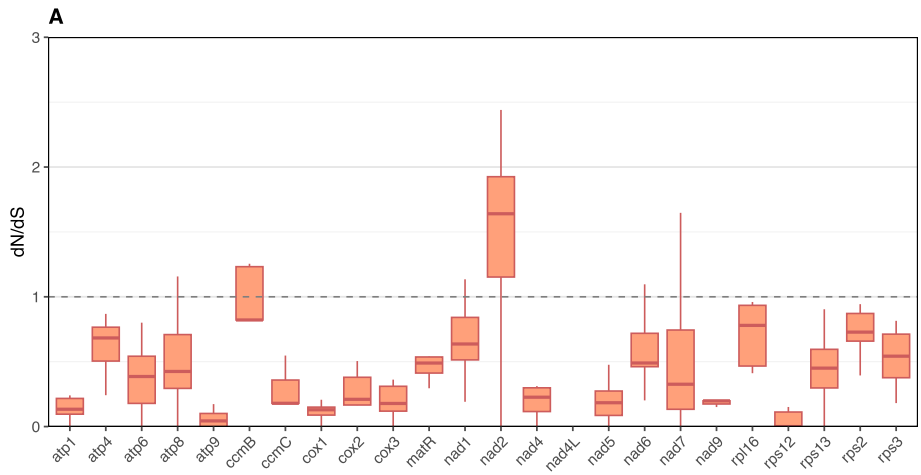


**Fig. 5** dN/dS analysis of PCGs in the mitogenome of 6 closely related species. The red solid lines on the box plot represents the average value

**Analysis of homologous fragments of MTPTs and NUMTs**

The analysis revealed a total of 30 homologous regions shared between mitochondrial and chloroplast genomes (MTPTs) of *S. sagittifolius*, with a combined length of 36,440 bp (Fig. 6A, Table S8). The longest fragment among these, MTPT4, spans 5,179 bp. The homologous fragments were annotated to include tRNA genes (*trnA-UGC, trnC-GCA, trnF-GAA, trnH-GUG, trnI-CAU, trnM-CAU, trnN-GUU, trnP-UGG, trnW-CCA* ) and plastid genes (*rpl2, rpl22, rpl23, ycf2, ycf3, ycf15,*

*rps3, rps11, rps19, rpoA, rpoB, rpoC1, rpoC2, psaA, psaB, psbA, psbB, petD, ycf3, atpA, ndhI, rbcL*). We identified 4,018 homologous regions shared between nuclear and mitochondrial genomes (NUMTs) of *S. sagittifolius*, totaling 1,815,279 bp, with the longest fragment measuring 36,441 bp (Fig. 6B, Table S9). Among these fragments, 28 exceed 10 kb in length, while the majority of mitochondrial fragments are shorter than 1 kb.

### Uncovering RNA editing events

RNA editing is a critical post-transcriptional process in plant mitochondria, primarily involving the conversion of cytidine (C) to uridine (U) in mRNA sequences [48]. An analysis of the lncRNA results from *S. sagittifolius* identified 776 RNA editing sites, of which 336 were associated with 27 protein-coding genes (PCGs). After filtering out false positives, 283 RNA editing sites were confirmed (Table S10), all identified sites underwent C-to-U editing. These edits predominantly occurred at the first and second positions of codons, with a higher frequency observed at the second position. As illustrated in Fig. 7A, among the core genes, nad1 exhibited the highest number of editing sites (23), followed by nad2 (22) and ccmFN (21), which are involved in photoprotection and energy metabolism. Mutation-type analysis revealed that synonymous mutations accounted for 49.65% (142 sites), nonsynonymous mutations for 44.06% (126 sites), and nonsense mutations for 6.29% (18 sites). Among the 144 nonsynonymous and nonsense mutations (Fig. 7B), 27 sites (18.75%) involved transitions from hydrophilic to hydrophobic amino acids, while 9 sites (6.25%) represented the reverse transition. The remaining 75.00% did not alter the physicochemical properties of the amino acids. Additionally, RNA editing showed a notable preference for codon transitions encoding lysine (25 sites) and leucine (22 sites) (Fig. 7B).

To assess the accuracy of predicted RNA editing sites, we randomly selected the *rpl16*, *atp6*, *atp9*, and *nad4L* genes for analysis. PCR amplification was conducted using both gDNA and cDNA as templates (Fig. 7C, Figure S6A, Table S11-S12), followed by Sanger sequencing (Fig. 7D, Figure S6B). For *rpl16*, six of the eight predicted RNA editing sites were confirmed, and two additional novel editing sites were detected. However, no RNA editing sites were identified for *atp6*, *atp9*, or *nad4L*.

### Phylogenomic and synteny analysis

Both mithchondrial and chloroplast genome analyses show a congruent phylogenetic structure that *S. sagittifolius* falls in the Andropogoneae tribe of Panicoideae subfamily and occupies a basal phylogenetic position relative to *C. zizanioides*, *C. lacryma-jobi* var. *maxima*, and *C. lacryma-jobi* var. *ma-yuen* (BS = 100; Fig. 8 A-B).

To assess the relationship between the mitochondrial genome of *S. sagittifolius* and its three closest relatives in this study, we identified conserved collinearity blocks consisting of homologous genes (Fig. 9). The analysis revealed significant variation in the arrangement of these collinear blocks among the mitochondrial genomes, suggesting extensive gene rearrangements. As illustrated in Figure S7, the mitogenomes of *S. sagittifolius*, *Zea luxurians*, *Chrysopogon zizanioides*, and *C. lacryma-jobi* var. *maxima* exhibit substantial structural divergence, further highlighting the rearrangements compared to closely related species.

## Discussion

### Mitochondrial structural stability and plasticity

The mitochondrial genome of *S. sagittifolius* (500,699 bp) is comparable in size to that of *Avena longiglumis* (548,445 bp). However, *A. longiglumis* possesses a mitogenome comprising four circular chromosomes [49], highlighting significant structural differences despite their similar sizes. Other species exhibit even greater structural diversity, such as *Punica granatum,* with seven circular contigs totaling 382,774 bp [50], and *Salvia officinalis*,which has two circular chromosomes of 268,341 bp and 39,827 bp [51]. The single circular mitochondrial genome of *S. sagittifolius* represents a relatively simple and conserved structural form, suggesting stable genomic organization that reflects its unique evolutionary trajectory and adaptation.

A noteworthy finding in this study is the absence of a positive correlation between genome size and repetitive sequence content in *S. sagittifolius*. This contradicts the prevailing view that mitogenome expansion is primarily driven by the accumulation of repetitive sequences [52].

(See figure on next page.)

**Fig. 6** Schematic representations of mitochondrial plastid DNAs (MTPTs) and nuclear mitochondrial DNA segments (NUMTs) in *S. sagittifolius*. **A** Comparative analysis of sequences from the mitochondrial genome and the chloroplast genome. The outer circle displays the mitogenome (mtDNA) in purple and the chloroplast (cpDNA) in green. Chloroplast genes are depicted in dark purple, mitochondrial genes in yellow with labeled gene names, and unannotated fragments in light purple. **B** Comparative analysis of sequences from the mitochondrial and nuclear genomes. The nuclear chromosomes are highlighted with a yellow square. Green arcs connect homogeneous sequence fragments between the mitochondrial and nuclear genomes

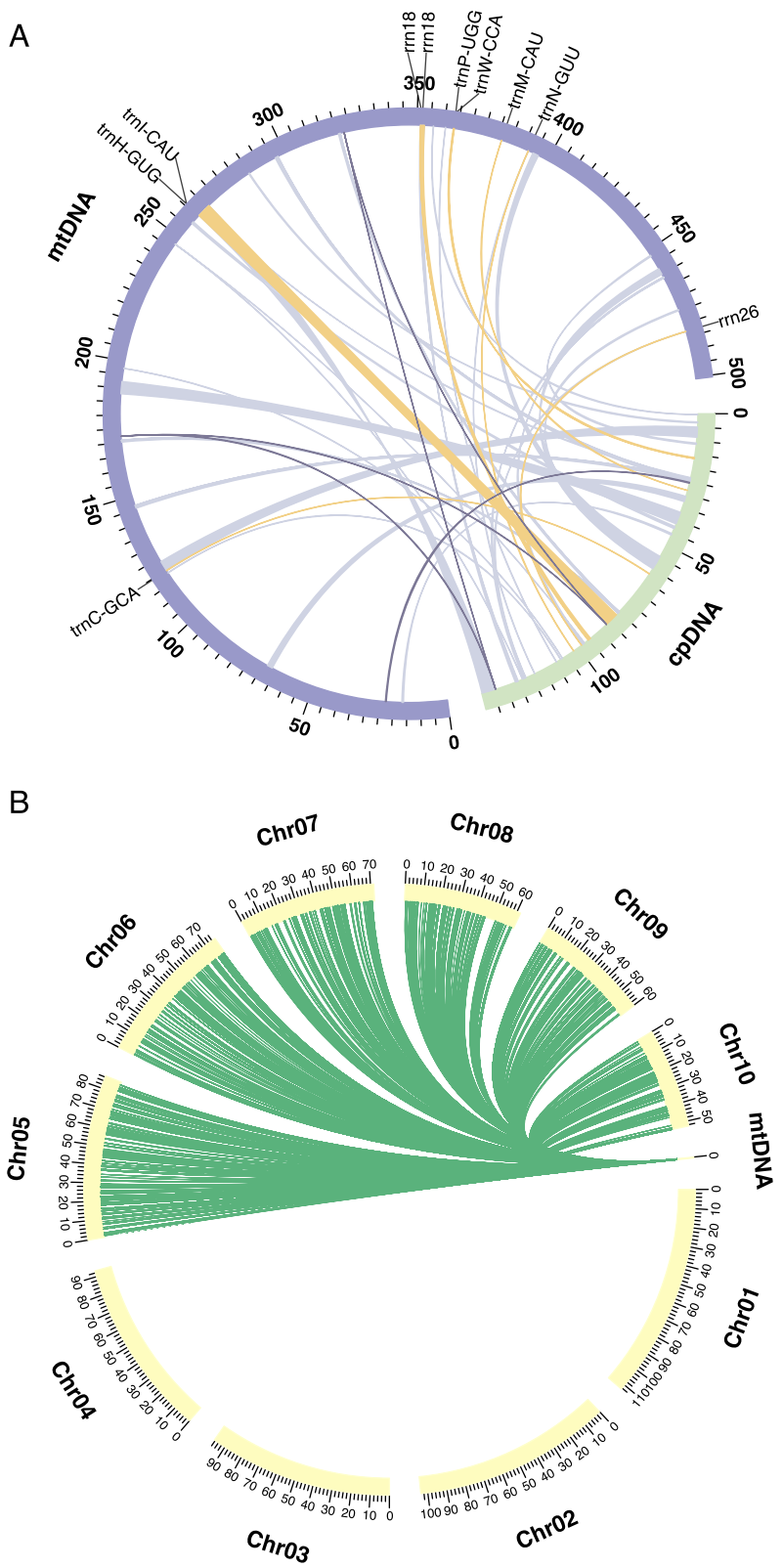


**Fig. 6** (See legend on previous page.)

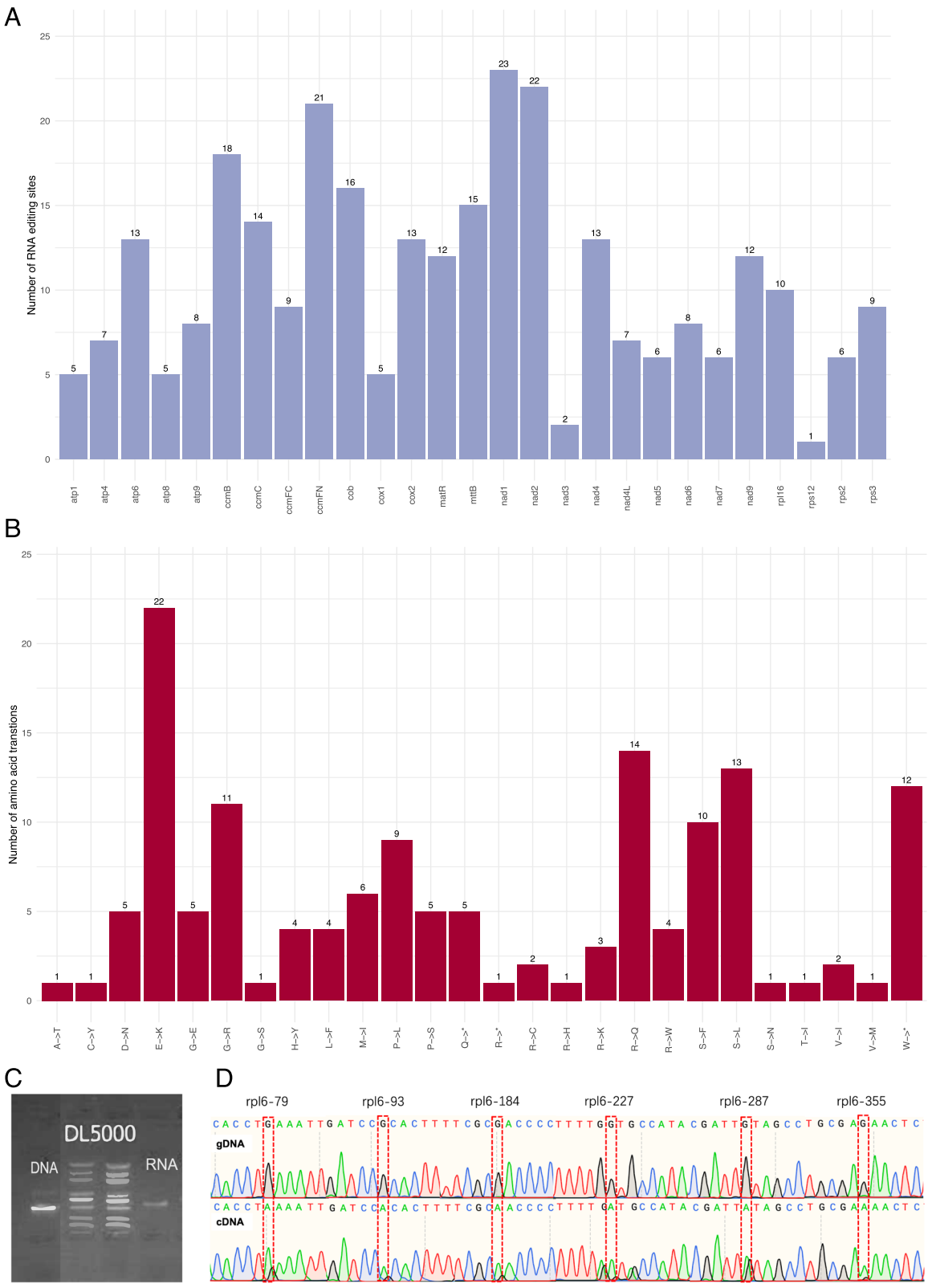


**Fig. 7** Prediction and validation of RNA editing sites in the mitogenomes of *S. sagittifolius*. (**A**) The number of RNA editing sites, (**B**) The number of amino acids with missense mutations, (**C**) PCR amplification product electrophoresis detection of the rpl16 gene, (**D**) Comparison of gDNA and cDNA editing sites of the *rpl16* gene coding region

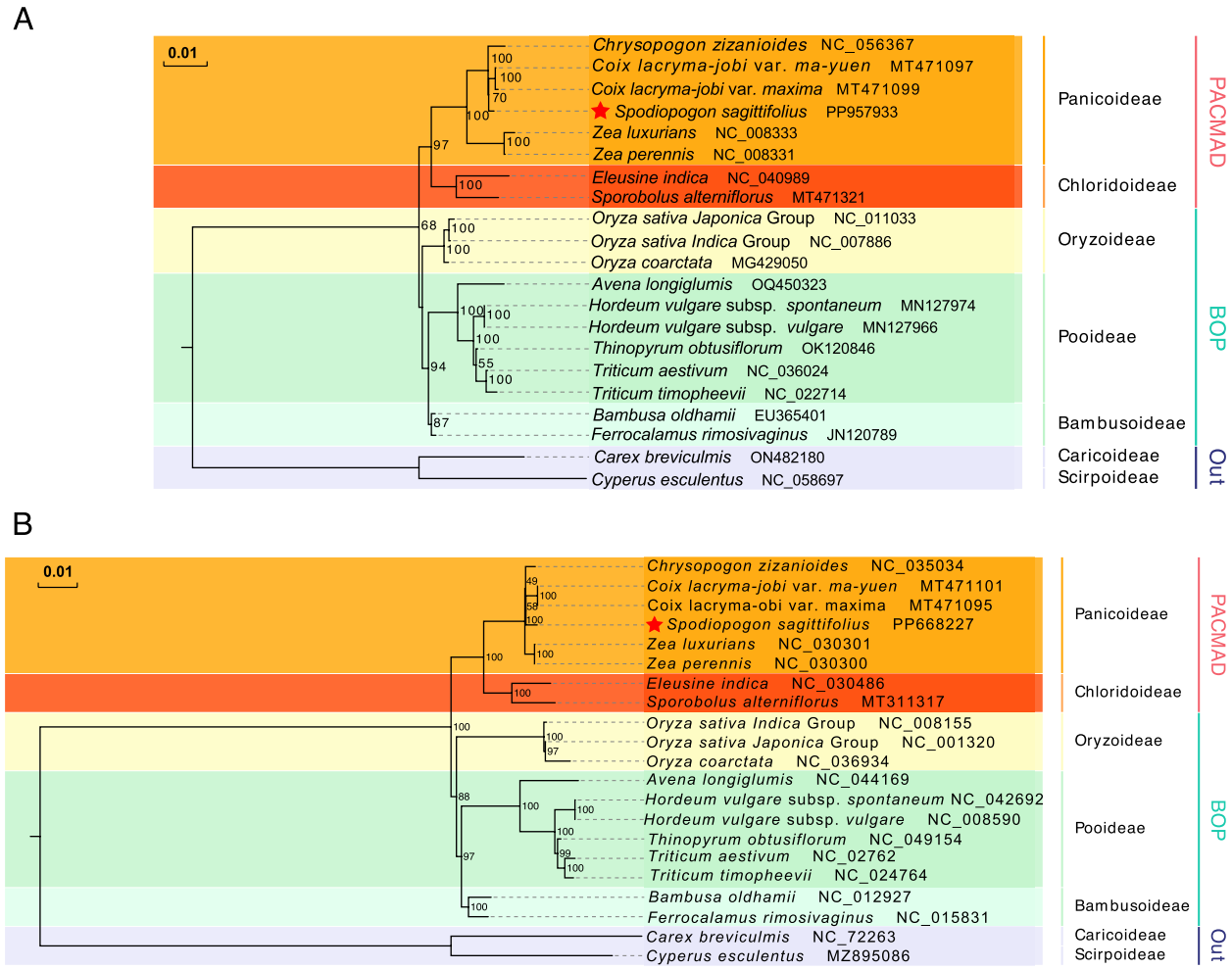


**Fig. 8** The phylogenetic trees between *S. sagittifolius* and 20 other species. (**A**) and (**B**) are phylogenetic trees constructed based on orthologous genes of mitochrial and chloroplast genomes respectively. Colors represent plants of the same subfamily. The bootstrap score was obtained using 1000 replicates. The mitochondrial and chloroplast genomes assembled in this study are labeled with five-pointed star

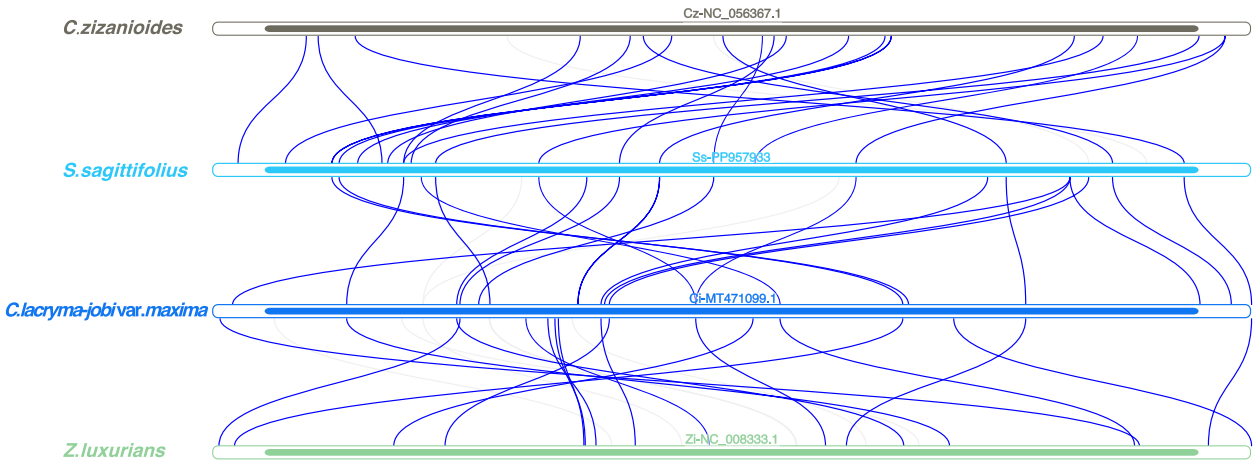


**Fig. 9** The collinear blocks between the mitogenomes of four closely related species in Panicoideae. The bar chart represents the mitogenome, and the arc represents the homologous sequence of adjacent four species. The blue highlighted area represents homologous genes with lengths greater than 500 bp, while the gray area represents homologous genes with lengths less than 500 bp

Instead, the expansion of plant mitochondrial genomes, as seen in *S. sagittifolius*, appears to be predominantly influenced by the insertion of non-coding regions and the integration of exogenous DNA, including fragments derived from nuclear or plastid genomes. Repetitive sequences in this context likely play a functional role in facilitating recombination events rather than directly affecting genome size. Furthermore, genome size is closely associated with gene loss [52]. This nuanced perspective challenges conventional assumptions and underscores the complex interplay between genome architecture and evolutionary forces.

The loss of multiple ribosomal protein genes (*rps1, rps14, rps19, rps4, rps7, rpl2, rpl23, rpl5*) in *S. sagittifolius* further emphasizes the dynamic evolution of its mitochondrial genome. Such gene loss may result from functional transfer to the nuclear genome, a well-documented phenomenon driven by the inherent instability and rearrangements of plant mitochondrial genomes. However, this study did not confirm whether homologous sequences corresponding to these genes have been transferred to the nuclear genome, leaving this hypothesis unverified. Alternatively, functional redundancy, where other subunits or protein complexes compensate for the roles of these missing genes, may explain their absence [51–53]. These possibilities highlight the evolutionary plasticity of plant mitochondrial genomes and underscore the distinct evolutionary trajectory of *S. sagittifolius*, even in comparison with closely related species.

The incorporation of mitochondrial DNA fragments into nuclear genomes (NUMTs) is a common and ongoing phenomenon in plant evolution. Gene transfer events often coincide with increased genetic drift, particularly during periods of environmental change or species adaptation [54]. While NUMTs contribute to genomic interactions and variation, they may also introduce risks, such as deleterious mutations in energy-demanding tissues [55–57]. In contrast, chloroplast genomes exhibit high conservation and minimal homologous recombination, providing valuable taxonomic insights [58]. Interestingly, our study revealed a greater number of homologous fragments between the mitochondrial and nuclear genomes than between the mitochondrial and chloroplast genomes. This observation reinforces the dynamic interactions between mitochondrial and nuclear genomes in shaping genomic diversity.

The unique genomic features observed in *S. sagittifolius* likely reflect species-specific evolutionary trajectories within plant mitochondrial genomes. Future studies should explore the functional significance of gene transfer events and investigate the mechanisms driving structural and functional variations in plant mitochondrial genomes. Such research will enhance our understanding of plant genome evolution and shed light on the adaptive strategies underlying genomic diversity across different ecological contexts.

### Putative adaptive signatures in the mitochondrial genome of S. sagittifolius

In our analysis, we found that the *sdh4* gene has been lost in 11 closely related species, except for *A. semialata* and *S. sagittifolius*. The *sdh4* gene plays a crucial role in the mitochondrial electron transport chain, specifically as a part of the succinate dehydrogenase (SDH) complex, also known as Complex II [59, 60]. This complex is involved in both the tricarboxylic acid (TCA) cycle and the mitochondrial respiratory chain, facilitating the conversion of succinate to fumarate and the transfer of electrons to the quinone pool [59]. The *sdh4* gene's presence provides a basis for research in mitochondrial function, evolutionary biology, stress tolerance, genomics, and crop improvement [60], in future research, this gene warrants significant attention.

The *nad2* gene encodes a subunit of NADH dehydrogenase (Complex I), a key component of the mitochondrial respiratory chain responsible for ATP synthesis through energy conversion [61]. Positive selection acting on *nad2* suggests its adaptive evolution to optimize energy production under diverse environmental stresses and varying metabolic demands during growth and development. These findings underscore the critical role of *nad2* in maintaining mitochondrial efficiency under dynamic biotic and abiotic conditions. In *S. sagittifolius*, which inhabits the dry-hot limestone regions of river valleys, positive selection on *nad2* likely enhances mitochondrial functionality, enabling efficient energy metabolism to sustain cellular energy balance under environmental stress. This adaptation is likely a crucial factor in improving the species' resilience to arid and high-temperature conditions, thereby supporting its growth and survival in such challenging habitats. Further investigations into the functional consequences of *nad2* positive selection in *S. sagittifolius* are essential to clarify its role in the plant's physiological adaptations and ecological strategies.

The mitochondrial genomes of *S. sagittifolius* and its closely related species (*C. zizanioides*, *C. lacryma-jobi* var. *maxima,* and *Z. luxurians*) exhibit limited structural rearrangements, suggesting a high degree of genomic stability among these taxa. This is supported by phylogenetic analyses, which place these species within the same clade, indicating a recent common ancestor [1]. This stability may reflect their shared evolutionary history, which have minimized the need for extensive genomic reorganization. The low frequency of rearrangements in closely related species is likely influenced by the absence of large repeat sequences that typically drive homologous

recombination. In *S. sagittifolius*, the mitochondrial genome contains fewer large repeats, which may explain its structural stability. This is consistent with findings in other plant groups, where repeat-mediated recombination is a major driver of mitochondrial genome plasticity [5].

### Exploring the adaptive role of RNA editing and codon bias

RNA editing is a key post-transcriptional modification in plant mitochondria, characterized predominantly by C-to-U conversions. This process is critical for correcting errors, refining protein-coding sequences, and regulating gene expression [62–64]. The extent and pattern of RNA editing vary substantially across species, tissues, and environmental conditions [65]. For example, only nine editing sites are present in the mtDNA of *Cyperus stoloniferus* [66], compared to 451 in *Salvia officinalis* [51], 546 in *Corydalis saxicola* [67] and 269 in *Avena longiglumis* [49]. In *S. sagittifolius*, although not all genes were validated, the absence of confirmation for some predicted RNA editing sites does not imply that the predictions are entirely inaccurate. RNA editing is a dynamic process influenced by cellular states and environmental factors. The patterns and efficiencies of RNA editing can vary across different cell types, developmental stages, or physiological conditions [66]. In this study, the absence of editing sites in genes such as *atp9*, *atp6*, and *nad4L* may be attributed to specific cellular conditions in leaf tissues at different developmental stages and environmental conditions, where RNA editing does not occur.

By altering codon composition, RNA editing optimizes mitochondrial protein compatibility with the translational machinery, enhancing their roles in energy metabolism and stress adaptation. This interplay between RNA editing and codon bias likely underpins the resilience of *S. sagittifolius* to its native environment dry-hot limestone regions of river valleys. The combined effects of these processes may support mitochondrial efficiency, ensuring survival and growth under environmental stress. Future studies should investigate the molecular mechanisms linking RNA editing and codon bias, with an emphasis on their collective influence on mitochondrial function and plant adaptation to diverse habitats [68]. These efforts will advance understanding of the evolutionary and ecological significance of post-transcriptional modifications in plant mitochondrial genomes.

## Conclusions

In this study, we successfully assembled and annotated the first mitochondrial genome of the genus *Spodiopogon* using the PMAT tool and HiFi sequencing data. The mitochondrial genome of *S. sagittifolius* is 500,699 bp in length with a GC content of 43.15%. The identification of repeat regions provides potential genetic markers for population genetics and evolutionary studies. Synteny and dN/dS analyses revealed the structural and functional conservation of mitochondrial genomes among closely related species. Most protein-coding genes exhibited low rearrangement frequencies and were under purifying selection (dN/dS < 1), indicating that the mitochondrial genome of *S. sagittifolius* is highly optimized for energy metabolism and cellular functions. However, the positive selection observed in the *nad2* gene (dN/dS = 1.49) suggests adaptive evolution, potentially in response to environmental stress. Extensive RNA editing events were identified across 27 protein-coding genes, primarily involving C-to-U conversions, with synonymous mutations accounting for 49.65% of editing sites. These editing events, coupled with a strong codon usage bias favoring A/U-ending codons, likely reflect the species' adaptation to its native environment, enhancing mitochondrial efficiency and stress tolerance. Phylogenetic analyses based on mitochondrial and chloroplast genomes confirmed the taxonomic placement of *S. sagittifolius* within the tribe Andropogoneae, providing new insights into its evolutionary relationships. In conclusion, this study advances our understanding of the evolutionary dynamics of the mitochondrial genome in *S. sagittifolius* and establishes a foundation for its conservation and potential crop domestication.

### Abbreviations

| | |
|---|---|
| BI | Bayesian Inference |
| BLAST | Basic local alignment search tool |
| cpDNA | Chloroplast DNA |
| dN/dS | Nonsynonymous to synonymous substitution rate ratio |
| GC | Guanine-cytosine |
| Mitogenome | Mitochondrial genome |
| mtDNA | Mitochondrial DNA |
| MTPTs | Regions shared between mitochondrial and chloroplast genomes |
| NUMTs | Regions shared between Nuclear and mitochondrial genomes |
| NCBI | National Center for Biotechnology Information |
| PCGs | Protein-coding genes |
| RSCU | Relative synonymous codon usage |
| rRNA | Ribosomal RNA |
| SSRs | Simple sequence repeats |
| TE | Tandem repeats |
| tRNA | Transfer RNA |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06341-z.

Additional file 1: Figure S1. Mitogenome assembly results of S. sagittifolius based on PMAT. Figure S2. Diagram depicting the precise positions of exons and introns involved in cis- and trans-splicing in S. sagittifolius. Figure S3. Depth of sequence coverage based on Illumina short-reads. Figure S4. The genes presence and absence of 14 species selected from Poaceae, A-N stands for Oryza sativa Japonica Group (NC_011033), Alloteropsis semialata (MH644808), Chrysopogon zizanioides (NC_056367), Coix lacryma-jobi var. maxima (MT471099), Coix lacryma-jobi var. ma-yuen

(MT471097), Coix lacryma-jobi var. puellarum (MT471098), Sorghum bicolor subsp. Drummondii (MZ506736), Sorghum bicolor (NC_008360), Spodiopogon sagittifolius (PP957933, in this study), Tripsacum dactyloides (NC_008362), Zea luxurians (NC_008333), Zea mays subsp. mays (NC_007982), Zea mays subsp. parviglumis (NC_008332), Zea perennis (NC_008331) respectively. A capital letter marked with an asterisk indicates the target species. Purple and gray rectangles denote for complete and missing mtDNA genes, respectively.Figure S5. Detected repeats in the S. sagittifolium mitogenomes. A & B Type and number of detected SSRs repeats. C Type and number of detected tandem and dispersed repeats. Figure S6. A) PCR amplification products of DNA and RNA for atp6, atp9, and nad4L genes. B) Comparison of gDNA and cDNA sequences of the rpl16 gene, with highlighted red sections indicating RNA editing sites. Figure S7. Dot plot analysis of the mitochondrial sequences

Additional file 2: Table S1. Characteristics of unitigs assembled by PMAT. Table S2. Annotated genes and their locations. Table S3. Number of protein-coding genes in mitochondrial genome of 14 species of Panicoideae. Table S4. SSRs identified in the mitochondrial genome of S. sagittifolius. Table S5. Dispersed repeats (≥ 30 bp) identified in the mitochondrial genome of S. sagittifolius.Table S6. Tandem repeat sequences identified in the mitochondrial genome of S. sagittifolius.Table S7. Selection pressure calculation in the mitochondrial protein coding genes of S. sagittifolius. Table S8. The homologous DNA fragment identified among the mitochondrial genome and plastome of S. sagittifolius. Table S9. The blastn results among the mitogenome (Query) and nuclear genome (Subject) of S. sagittifolius. Table S10. RNA editing events identified in the mitochondrial protein coding genes of S. sagittifolius. Table S11. PCR primers for verifying RNA editing sites of rpl16 gene of S. sagittifolius. Table S12. Species and GenBank accessions used for phylogenetic analysis.

**Acknowledgements**
We sincerely thank Dr. Yanbin Guo from Yunnan Agricultural University for his guidance and suggestions on mitochondrial analysis in this study.

**Clinical trial number**
Not applicable.

**Authors' contributions**
CX and ZWL conceived the research. ZWL assembled the mitogenome, carried out the analysis of the phylogeny and prepared software. CX performed the bioinformatic analyses. CX and WB wrote the paper. RYM prepared figures and tables. FL and PRL collected and prepared the samples and revised the manuscript.

**Data availability**
No datasets were generated or analysed during the current study.

## Declarations

**Competing interest**
The authors declare no competing interests.

Received: 6 September 2024 Accepted: 4 March 2025
Published online: 24 March 2025

**References**

1. Soreng RJ, Peterson PM, Zuloaga FO, Romaschenko K, Clark LG, Teisher JK, et al. A worldwide phylogenetic classification of the Poaceae (Gramineae) III: An update. J Syst Evol. 2022;60:476–521.
2. Armstrong EM, Larson ER, Harper H, Webb CR, Dohleman F, Araya Y, et al. One hundred important questions facing plant science: an international perspective. New Phytol. 2023;238:470–81.
3. Jin Z. Characteristics and Utilization of Shrub and Grass Vegetation in Tropical and Subtropical Mountainous Areas of Yunnan. Acta Phytoecologica Sinica. 1986;10:81.
4. Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, et al. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 2012;10:e1001241.
5. Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, Michelmore RW, et al. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. PLoS Genet. 2019;15:e1008373.
6. Mower JP, Sloan DB, Alverson AJ. Plant mitochondrial genome diversity: the genomics revolution. Plant genome diversity volume 1: plant genomes, their residents, and their evolutionary dynamics. 2012;1:123–44.
7. Van der Bliek AM, Shen Q, Kawajiri S. Mechanisms of mitochondrial fission and fusion. Cold Spring Harb Perspect Biol. 2013;5:a011072.
8. Palmer JD, Herbon LA. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol. 1988;28:87–97.
9. Hiesel R, Von Haeseler A, Brennicke A. Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. Proc Natl Acad Sci USA. 1994;91:634–8.
10. Gualberto JM, Newton KJ. Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. Annual Review of Plant Biology. 2017;68 Volume 68, 2017:225–52.
11. Schertl P, Braun H-P. Respiratory electron transfer pathways in plant mitochondria. Front Plant Sci. 2014;5:163.
12. Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature. 2018;557:43–9.
13. Exposito-Alonso M, Burbano HA, Bossdorf O, Nielsen R, Weigel D. Natural selection on the Arabidopsis thaliana genome in present and future climates. Nature. 2019;573:126–9.
14. Arseneau J-R, Steeves R, Laflamme M. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. Mol Ecol Resour. 2017;17:686–93.
15. Bi C, Qu Y, Hou J, Wu K, Ye N, Yin T. Deciphering the Multi-Chromosomal Mitochondrial Genome of Populus simonii. Frontiers in Plant Science. 2022;13:914635.
16. Bi C, Shen F, Han F, Qu Y, Hou J, Xu K, et al. PMAT: an efficient plant mitogenome assembly toolkit using low-coverage HiFi sequencing data. Horticulture Research. 2024;11:uhae023.
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
18. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 2015;31:3350–2.
19. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997. 2013.
20. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE. 2014;9:e112963.
21. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021;10:giab008.
22. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45:W6-11.
23. Lewis SE, Searle SMJ, Harris N, Gibson M, Lyer V, Richter J, et al. Apollo: a sequence annotation editor. Genome Biol. 2002;3:RESEARCH0082.
24. Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, et al. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 2019;47:W65-73.
25. Chen C, Wu Y, Li J, Wang X, Zeng Z, Xu J, et al. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. Mol Plant. 2023;16:1733–42.
26. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.
27. Picardi E, Pesole G. REDItools: high-throughput RNA editing detection made easy. Bioinformatics. 2013;29:1813–4.
28. Robinson JT, Thorvaldsdottir H, Turner D, Mesirov JP. igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). Bioinformatics. 2022;39:btac830.
29. Weber S, Ramirez C, Doerfler W. Signal hotspot mutations in SARS-CoV-2 genomes evolve as the virus spreads and actively replicates in different parts of the world. Virus Res. 2020;289:198170.

30. Xiang C, Gao F, Jakovlić I, Lei H, Hu Y, Zhang H, et al. Using PhyloSuite for molecular phylogeny and tree-based analyses. iMeta. 2023;2:e87.
31. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 2020;20:348–55.
32. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular biology and evolution. 2018;35:1547–9.
33. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33:2583–5.
34. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29:4633–42.
35. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–80.
36. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
37. Null R Core Team R, Team R, Null R Core Team, Core Writing T, Null R, Team R, et al. R: A language and environment for statistical computing. Computing. 2011;1:12–21.
38. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
39. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540–52.
40. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37:1530–4.
41. Krumsiek J, Arnold R, Rattei T. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics. 2007;23:1026–8.
42. Yang Z, Nielsen R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol Biol Evol. 2000;17:32–43.
43. Rashmi D, Bisen P, Saha S, Loitongbam B, Singh S, Singh P, et al. Genetic diversity analysis in rice (Oryza sativa L) accessions using SSR markers. International Journal of Agriculture, Environment and Biotechnology. 2017;10:457–67.
44. Mishra A, Chaudhary S, Kumar A, Chandrasekharan H. Identification of Simple Sequence Repeats in chloroplast and mitochondrial genome of wheat. In: 2014 International Conference on Computing for Sustainable Global Development (INDIACom). IEEE; 2014:265–70.
45. Li JQ, Jiong F. Genomic diversity and evolutionary mechanisms in the Oryza genus: a comparative analysis. Genomics Appl Biol. 2024;15(1):54–63.
46. Christensen AC. Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. Genome Biol Evol. 2013;5:1079–86.
47. Galtier N. The intriguing evolutionary dynamics of plant mitochondrial DNA. BMC Biol. 2011;9:1–3.
48. Bruhs A, Kempken F. RNA Editing in Higher Plant Mitochondria. In: Kempken F, editor. Plant Mitochondria. New York: Springer New York; 2011. p. 157–175.
49. Liu Q, Yuan H, Xu J, Cui D, Xiong G, Schwarzacher T, et al. The mitochondrial genome of the diploid oat Avena longiglumis. BMC Plant Biol. 2023;23:218.
50. Feng L, Wang Z, Wang C, Yang X, An M, Yin Y. Multichromosomal mitochondrial genome of Punica granatum: comparative evolutionary analysis and gene transformation from chloroplast genomes. BMC Plant Biol. 2023;23:512.
51. Yang H, Chen H, Ni Y, Li J, Cai Y, Wang J, et al. Mitochondrial genome sequence of Salvia officinalis (Lamiales: Lamiaceae) suggests diverse genome structures in cogeneric species and finds the stop gain of genes through RNA editing events. Int J Mol Sci. 2023;24:5372.
52. Yang H, Ni Y, Zhang X, Li J, Chen H, Liu C. The mitochondrial genomes of Panax notoginseng reveal recombination mediated by repeats associated with DNA replication. Int J Biol Macromol. 2023;252:126359.
53. Figueroa P, Léon G, Elorza A, Holuigue L, Araya A, Jordana X. The four subunits of mitochondrial respiratory complex II are encoded by multiple nuclear genes and targeted to mitochondria in Arabidopsis thaliana. Plant Mol Biol. 2002;50:725–34.
54. Butenko A, Lukeš J, Speijer D, Wideman JG. Mitochondrial genomes revisited: why do different lineages retain different genes? BMC Biol. 2024;22:15.
55. Liu S-L, Zhuang Y, Zhang P, Adams KL. Comparative Analysis of Structural Diversity and Sequence Evolution in Plant Mitochondrial Genes Transferred to the Nucleus. Mol Biol Evol. 2009;26:875–91.
56. Puertas MJ, González-Sánchez M. Insertions of mitochondrial DNA into the nucleus—effects and role in cell evolution. Genome. 2020;63:365–74.
57. Lim K. Mitochondrial genome editing: strategies, challenges, and applications. BMB Rep. 2024;57:19.
58. Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V, et al. Angiosperm phylogeny based on <011>matK sequence information. Am J Bot. 2003;90:1758–76.
59. Li Y, Belt K, Alqahtani SF, Saha S, Fenske R, Van Aken O, et al. The mitochondrial LYR protein SDHAF1 is required for succinate dehydrogenase activity in Arabidopsis. Plant J. 2022;110:499–512.
60. Huang S, Braun H-P, Gawryluk RM, Millar AH. Mitochondrial complex II of plants: subunit composition, assembly, and function in respiration and signaling. Plant J. 2019;98:405–17.
61. Milenkovic D, Blaza JN, Larsson N-G, Hirst J. The enigma of the respiratory chain supercomplex. Cell Metab. 2017;25:765–76.
62. Chateigner-Boutin A-L, Small I. Plant RNA editing. RNA Biol. 2010;7:213–9.
63. Takenaka M, Zehrmann A, Verbitskiy D, Härtel B, Brennicke A. RNA editing in plants and its evolution. Annu Rev Genet. 2013;47:335–52.
64. Shikanai T. RNA editing in plant organelles: machinery, physiological function and evolution. Cellular and Molecular Life Sciences CMLS. 2006;63:698–708.
65. Hammani K, Giegé P. RNA metabolism in plant mitochondria. Trends Plant Sci. 2014;19:380–9.
66. Miao X, Yang W, Li D, Wang A, Li J, Deng X, et al. Assembly and comparative analysis of the complete mitochondrial and chloroplast genome of Cyperus stoloniferus (Cyperaceae), a coastal plant possessing saline-alkali tolerance. BMC Plant Biol. 2024;24:1–18.
67. Li C, Liu H, Qin M, Tan Y, Ou X, Chen X, et al. RNA editing events and expression profiles of mitochondrial protein-coding genes in the endemic and endangered medicinal plant. Corydalis saxicola Frontiers in Plant Science. 2024;15:1332460.
68. Tang W, Luo C. Molecular and functional diversity of RNA editing in plant mitochondria. Mol Biotechnol. 2018;60:935–45.


## Publisher's Note