frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 August 2026
DOI 10.3389/fpls.2026.1926603

Check for updates

OPEN ACCESS

EDITED BY
Cun Guo,
Tarim University, China

REVIEWED BY
Jiangui Zhang,
Yunnan Agricultural University, China
Xiqiang Liu,
Chinese Academy of Forestry, China

*CORRESPONDENCE
Xiaoming Bai
baixm@gsau.edu.cn
Zinian Wu
wuzinian@caas.cn

RECEIVED 02 July 2026
REVISED 23 July 2026
ACCEPTED 12 August 2026
PUBLISHED 26 August 2026

CITATION
Teng Z, Tian C, Yang Y, Gong W, Li Z, Bai X and Wu Z (2026) Mitogenome assembly and phylogenetic relationships of *Phalaris arundinacea*. *Front. Plant Sci.* 17:1926603. doi: 10.3389/fpls.2026.1926603

COPYRIGHT
© 2026 Teng, Tian, Yang, Gong, Li, Bai and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Mitogenome assembly and phylogenetic relationships of *Phalaris arundinacea*

Ze Teng[1,2,3], Chunyu Tian[2,3], Yanting Yang[2,3], Wenlong Gong[2,3], Zhiyong Li[2,3], Xiaoming Bai[1]* and Zinian Wu[2,3]*

[1]Pratacultural College, Gansu Agricultural University, Lanzhou, China, [2]Institute of Grassland Research, Chinese Academy of Agricultural Sciences, Hohhot, China, [3]Key Laboratory of Grassland Resources and Utilization of Ministry of Agriculture, Hohhot, China

Introduction: As a perennial herb of Poaceae, *Phalaris arundinacea* plays key roles in grazing, production, and soil and water conservation because of its well-developed rhizomes and seed dispersal. We assembled and annotated the first mitogenome of *P. arundinacea* to support evolutionary and taxonomic research.
Methods: We assembled and annotated the first complete mitochondrial genome of *P. arundinacea* by integrating Illumina short reads with Nanopore long reads via a hybrid assembly strategy. The genome architecture was comprehensively characterized, encompassing codon usage bias, repetitive sequence organization, and inter-organellar genetic exchange with the chloroplast genome.
Results and discussion: Assembly of the *P. arundinacea* mitogenome revealed two circular structures with a combined length of 526,717 bp. The genome comprised a set of 37 protein-coding genes (PCGs), 27 tRNAs, and 8 rRNAs, with the rRNA genes exhibiting full assembly (100% coverage). The mitochondrial genome contained 154 forward and 164 palindromic repeats, along with 25 tandem repeats and 124 simple sequence repeats (SSRs). Notably, 102 SSRs were distributed on contig1, predominantly in tetrameric form. Furthermore, 376 RNA editing sites were predicted. A total of 104 fragments were integrated into the mitochondrial genome from the chloroplast, amounting to 55,866 bp of transferred sequence. Finally, phylogenetic analysis of 28 plant mitogenomes placed *P. arundinacea* closest to species within the genus *Poa* (*P. chaixii* and *P. pratensis*). Comparative analysis of non-synonymous-to-synonymous substitution rate (Ka/Ks) ratios across divergent species revealed that the mitochondrial genome of *P. arundinacea* underwent stabilizing evolutionary dynamics, characterized by predominant purifying selection with several lineage-specific variations in selective pressure. Our findings support the close phylogenetic relationship between P. arundinacea and species of the genus Poa and provide a reference mitochondrial genome resource for future comparative studies within Phalaris that incorporate broader taxon sampling. These results support deeper phylogenetic investigations of *P. arundinacea* and facilitate future work on its germplasm characterization and applied use.

KEYWORDS
gene transfer, mitogenome, *Phalaris arundinacea*, RNA editing, RSCU

Introduction

As a perennial herb of the Gramineae family, *Phalaris arundinacea* exhibits unique biological characteristics and a wide distribution range. The developed rhizome system, seed dispersal ability, and wide adaptability not only make it useful for grazing, production, and soil and water conservation, but also attract extensive attention as a bioenergy material. Owing to its strong adaptability, *P. arundinacea* thrives on marginal lands unsuitable for conventional agriculture and is naturally distributed across temperate and Neotropical regions (Sahramaa et al., 2004; Perdereau et al., 2017; Xiong et al., 2020). Meanwhile, *P. arundinacea* is regarded as an invasive species in some areas, competing with local crops for resources such as light, fertilizer, and water throughout its growth cycle (Apfelbaum and Sams, 1987; Lavergne et al., 2010). *P. arundinacea* is a polyploid plant with extensive variations in ploidy level, habitat, and geographical distribution. Initial molecular phylogenetic analyses assigned *P. arundinacea* to the Aveneae-Poeae complex within the subfamily Pooideae (Quintanar et al., 2007; Schneider et al., 2009; Saarela et al., 2015). Currently, 20 species of *Phalaris* have been identified (Voshell and Hilu, 2014). This includes both endemic species, such as *P. rotgesii*, *P. truncata*, *P. californica*, and *P. maderensis*, and the widespread cosmopolitan species *P. arundinacea* L (Galatowitsch et al., 1999; Lavergne and Molofsky, 2004; Thomsen et al., 2012).

As key cellular organelles, mitochondria and chloroplasts perform vital functions in plants, ranging from energy conversion and metabolism to developmental signaling (Chen and Butow, 2005). Compared with whole-genome sequencing, organellar genomes offer greater cost-effectiveness and serve as robust foundations for evolutionary studies (Sun et al., 2023). Plant mitochondrial genomes are characterized by considerable structural diversity—existing in both circular and linear forms—and are shaped by frequent repeat-mediated recombination, the acquisition of exogenous DNA, and the deletion of endogenous sequences (Bergthorsson et al., 2003; Sloan, 2013; Wynn and Christensen, 2019). Notably, plant mitochondrial genomes evolve rapidly in structure while slowly in sequence (Palmer and Herbon, 1988), making them particularly suitable for resolving structural rearrangements and inter-organellar DNA transfer events that are often undetectable using chloroplast or nuclear markers alone; Furthermore, plant cells frequently exhibit bidirectional DNA exchange between their chloroplast and mitochondrial genomes (Allen, 2015; Turmel et al., 2016). Collectively, these features make the mitochondrial genome a powerful tool for investigating structural variation, intracellular gene transfer, and phylogenetic relationships within plant groups (Zou et al., 2025).

In *P. arundinacea*, previous genomic research has achieved partial progress. The chloroplast genome has been utilized for species and ploidy comparisons (Xiong et al., 2020), and nuclear genome-based molecular markers such as EST-SSR and ISSR have revealed high genetic diversity within populations. However, these markers have limitations in elucidating the molecular mechanisms underlying complex adaptive traits. Crucially, a significant knowledge gap remains regarding the sequence architecture, structural configuration, and evolutionary trajectory of the *P. arundinacea* mitochondrial genome. Given its complex ecological adaptability and polyploid nature, key genes within the mitochondrial genome may regulate adaptive responses to diverse environments. Moreover, owing to the evident variations in its morphological characteristics and the crucial role that polyploidy plays in its evolutionary process (such as tetraploid and hexaploid *P. arundinacea*), the mitochondrial genome has the potential to provide novel insights into grass species formation, chromosome evolution, and biogeography (Xiong et al., 2020).

To address this gap, we assembled and annotated the first complete mitochondrial genome of P. arundinacea using a hybrid strategy integrating Illumina short reads with Nanopore long reads. Subsequently, with the help of these two platforms, we carried out detailed annotation of the assembled mitochondrial genome. Subsequent analyses encompassed multiple facets of the *P. arundinacea* mitogenomic architecture, including codon usage bias, repetitive sequence organization, and inter-organellar genetic exchange with the chloroplast genome. By constructing and analyzing a phylogenetic tree, we identified crucial clues to uncover the evolutionary history of *P. arundinacea*.

Materials and methods

DNA extraction, genome sequencing and assembly

Fresh leaf tissues of *P. arundinacea* were harvested from the National Perennial Forage Germplasm Resource Nursery located in Hohhot, Inner Mongolia, China (40.57°N, 111.93°E). To isolate high-quality genomic DNA, a modified protocol integrating the conventional CTAB method with the Qiagen Blood and Cell Culture DNA Kit (Cat. no. 13323) was implemented, followed by an ethanol precipitation step. Dual-platform sequencing was conducted using the Illumina NovaSeq 6000 and Oxford Nanopore PromethION systems. This combined approach yielded exactly 52,328,441 short reads (15,796,442,260 bases) from the Illumina platform and 1,578,326 long reads (17,189,341,717 bases) via Nanopore sequencing. For mitogenome reconstruction, the long Nanopore reads were initially aligned against a customized plant reference gene database (https://github.com/xul962464/plant_mt_ref_gene) utilizing minimap2 v2.1 (Sahlin, 2023). The resulting draft assembly underwent an error-correction polishing phase using Canu v2.2 (Koren et al., 2017). Subsequently, this polished sequence functioned as a reference framework to align the short Illumina reads using Bowtie2 v2.3.5.1 (Langmead and Salzberg, 2012). A comprehensive hybrid assembly was then constructed with Unicycler v0.4.8 (Wick et al., 2017) operating under standard default configurations. Finally, the structural integrity and circular conformation of the assembled mitogenome were visually validated and manually curated through the Bandage v0.8.1 graphical interface (Wick et al., 2015).

Genome annotation

For the initial annotation of the mitochondrial genome, we employed the GeSeq web server (Tillich et al., 2017), utilizing existing homologous Poaceae mitogenome records from the NCBI database as reference frameworks. The gene structure was subsequently predicted using PMGA (Li et al., 2025). The annotation results from both GeSeq and PMGA were integrated to generate a comprehensive and refined gene annotation profile, and Geneious (v2024.0.5) (Kearse et al., 2012) was employed for the manual curation and refinement of the initial annotation results. Finally, the genomic map was constructed using the OGDRAW (Organellar Genome Draw) software (v1.3.1) (Greiner et al., 2019).

Repeat sequence identification

Microsatellites (SSRs) within the *P. arundinacea* mitochondrial genome were mined using the MISA tool (Beier et al., 2017). The detection thresholds were rigorously standardized to a minimum of 10, 5, and 4 repeat units for mono-, di-, and trinucleotides, respectively, whereas a strict minimum of three iterations was required for tetra-, penta-, and hexanucleotide motifs. For tandem repeat identification, the FASTA-formatted assembly was processed using Tandem Repeats Finder (TRF v4.09) (Benson, 1999). The algorithm's scoring matrix was customized as follows: alignment match = 2, mismatch/indel penalties = 7, match probability (PM) = 80, and indel probability (PI) = 10. Additional constraints included a minimum alignment score of 50 and a maximum period size of 2,000 bp, alongside the activation of the -f (flanking), -d (data), and -m (masked) output parameters. Furthermore, REPuter (Kurtz et al., 2001) was executed to locate four distinct classes of dispersed repeats (forward, palindromic, reverse, and complement). The search constraints were restricted to a minimum length of 12 bp and a maximum Hamming distance of 3. Following stringency filtering, the surviving high-confidence repetitive elements were fully cataloged based on their specific structural classifications, genomic coordinates, and sizes.

Codon usage bias analysis

The relative synonymous codon usage (RSCU) profiles and amino acid frequencies for the entire set of PCGs were calculated using the CodonW software package (v1.4.4) (Iriarte et al., 2021). Codon usage bias was analyzed using custom Perl scripts, and unique coding sequences (CDS) were selected for subsequent analysis. Subsequently, the results were plotted using R.

Prediction of RNA editing sites

Potential RNA editing events within the mitochondrial protein-coding genes were identified using unprocessed sequencing data retrieved from SRA database (accession numbers: SRR13516133, SRR13516134, SRR13516135, SRR13516136, SRR13516137, SRR13516138, SRR13516139, SRR13516140, SRR13516141, SRR13516142, SRR13516143, and SRR13516144); https://www.ncbi.nlm.nih.gov/sra/). The alignment of RNA-seq reads to mitochondrial PCGs was executed using BWA v0.7.1 (Li and Durbin, 2010). Subsequently, variant calling was carried out via the SAMtools (v1.17) (Li et al., 2009) and BCFtools (v1.17) (Danecek et al., 2021) pipeline. To identify high-confidence RNA editing events, the resulting SNP profiles were systematically screened and annotated using the REDO toolkit (v1.0) (Wu et al., 2018). To reduce false positives, the DNA sequencing reads were mapped against the *Phalaris arundinacea* mitochondrial genome as the reference. Genomic SNPs were called using BCFtools (v1.17), followed by the exclusion of regions overlapping mRNA editing sites.

Chloroplast-derived mitochondrial sequence identification

Utilizing the same raw Illumina sequencing library, the plastid genome (cpDNA) was reconstructed through a combined protocol of *de novo* assembly and reference-guided alignment. The sequence assembly was executed using the GetOrganelle (v1.7.7.0) (Jin, 2020) under predefined configurations (-R 50 -k 21,45,65,85,105,115 -P 1000000). For gene prediction, an integrated approach deploying both CPGAVAS2 and PGA was applied (Qu et al., 2019; Shi et al., 2019). The preliminary automated outputs were subsequently merged and subjected to manual refinement within Geneious software to establish the definitive cpDNA annotation. To detect potential intracellular DNA transfers, the assembled chloroplast sequence was compared against the mitogenome via a homology search using BLASTn (v2.16.0) (Chen et al., 2015) with a strict cutoff parameter (E-value ≤ 1e-5). Finally, the spatial distribution and mapping of these plastid-derived fragments across the mitochondrial architecture were graphically represented using TBtools (v2.091) (Chen et al., 2020).

Selection pressure analysis and nucleotide diversity assessment

Sequence alignments for 27 conserved mitochondrial genes shared among *P. arundinacea* and 25 other Poaceae species were first generated using MAFFT (v7.526) (Katoh and Standley, 2013). These genes included *atp1, atp4, atp6, atp8, atp9, ccmB, ccmC, ccmFC, ccmFN, cox1, cox2, cox3, matR, mttB, nad1, nad2, nad3, nad4, nad4L, nad5, nad7, nad9, rpl16, rps13, rps2, rps3, and sdh4.* Based on the aligned sequences, the nonsynonymous (Ka) and synonymous (Ks) substitution rates were subsequently estimated via KaKs (Calculator v3.0) (Zhang, 2022). Additionally, the nucleotide diversity (Pi) across 30 specific mitochondrial genes in *P. arundinacea*, *Poa pratensis*, and *Poa chaixii* was quantified using DnaSP (v6.12.0) (Rozas et al., 2017).

Phylogenetic and synteny analysis

To elucidate the phylogenetic relationships, a matrix of 25 shared mitochondrial PCGs (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFC*, *ccmFN*, *cox1*, *cox2*, *cox3*, *matR*, *mttB*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad7*, *nad9*, *rpl16*, *rps3*, *sdh4*) was assembled from 26 Poaceae species, with *Arabidopsis thaliana* and *Medicago truncatula* included as outgroup taxa. All reference mitochondrial genome sequences were retrieved from the NCBI

repository (detailed in Supplementary Table 1). The selected PCGs were concatenated and aligned using MAFFT (v7.131) (Katoh and Standley, 2013). Following model selection via ModelFinder (Kalyaanamoorthy et al., 2017), phylogenetic inference was performed using two distinct approaches. The Maximum Likelihood (ML) tree was generated by RAxML (Stamatakis, 2006) employing the GTRGAMMA model with 1,000 bootstrap replicates, while Bayesian inference (BI) was executed in MrBayes (Huelsenbeck and Ronquist, 2001) under the GTR+F+I+G4 substitution model.

Synteny analysis

To elucidate genomic collinearity, pairwise sequence alignments were conducted between *P. arundinacea* and three related species (*Poa chaixii*, *Poa pratensis*, and *Triticum aestivum*) utilizing BLASTn (v2.16.0). Subsequently, AliTV (v1.0.6) (Ankenbrand et al., 2017) was employed to detect high-confidence syntenic blocks and construct comparative maps illustrating conserved genomic regions.

Results

Genomic features of the *P. arundinacea* mitogenome

The assembled *P. arundinacea* mitogenome comprises two complex, circular structures, spanning a total length of 526,717 bp with a GC content of 44.46% (Figures 1, 2). Annotation identified a total of 72 functional genes (Supplementary Tables 2, S3), which are unevenly distributed between contig1(ctg1) (58 genes) and contig2 (ctg2) (14 genes). The coding repertoire consists of 37 unique PCGs, 27 tRNAs, and 8 rRNAs. Among the PCGs, 16 are classified as core genes, including six ATP synthase subunits (with two copies of *atp1*, plus *atp4, atp6, atp8*, and *atp9*), five cytochrome c biogenesis genes (*ccmC, ccmFc, ccmFn*, and duplicated *ccmB*), *cob*, *cox1–3*, and the maturase *matR*. The remaining 21 PCGs are categorized as variable genes, encompassing the transport protein *mttB*, ten NADH dehydrogenase subunits (including duplicated *nad4L* and *nad1-7*, *nad9*), *rpl16*, and notably two copies of *rps19* alongside *rps1, rps3, rps4*, and *rps7, rps12-14*.

The mitochondrial genome assembly of *P. arundinacea* comprised two contigs: ctg1 (418,058 bp, GC content 44.62%) and ctg2 (108,659 bp, GC content 43.85%). ctg1 contained 31 CDS spanning 23,972 bp (GC 42.81%), 21 tRNA genes (1,785 bp, GC 50.14%), and eight rRNA genes (13,242 bp, GC 53.26%). In contrast, ctg2 harbored only six CDS (8,104 bp, GC 42.23%), six tRNA genes (462 bp, GC 53.25%), and no rRNA genes. Notably, the tRNA and rRNA regions exhibited higher GC content than the CDS in both contigs, suggesting distinct evolutionary constraints on non-coding functional elements.

Detailed annotation of gene structure (Supplementary Table 3) revealed distinct intron patterns within the *P. arundinacea* mitogenome. Specifically, the genes *nad1, nad2, nad5*, and *nad7* exhibited the most complex structure, each containing four introns, while *nad4* harbored three. *cox2* was found to contain two introns, whereas *ccmFc, rps3, trnA-TGC*, and *trnI-GAT* each possessed a single intron. Gene duplications were predominantly observed among non-coding RNAs. Three copies were identified for *rrn18, rrn5*, and *trnM-CAT*, while two copies were detected for *rrn26* and several tRNA genes (*trnD-GTC, trnE-TTC, trnP-TGG, trnQ-TTG,*

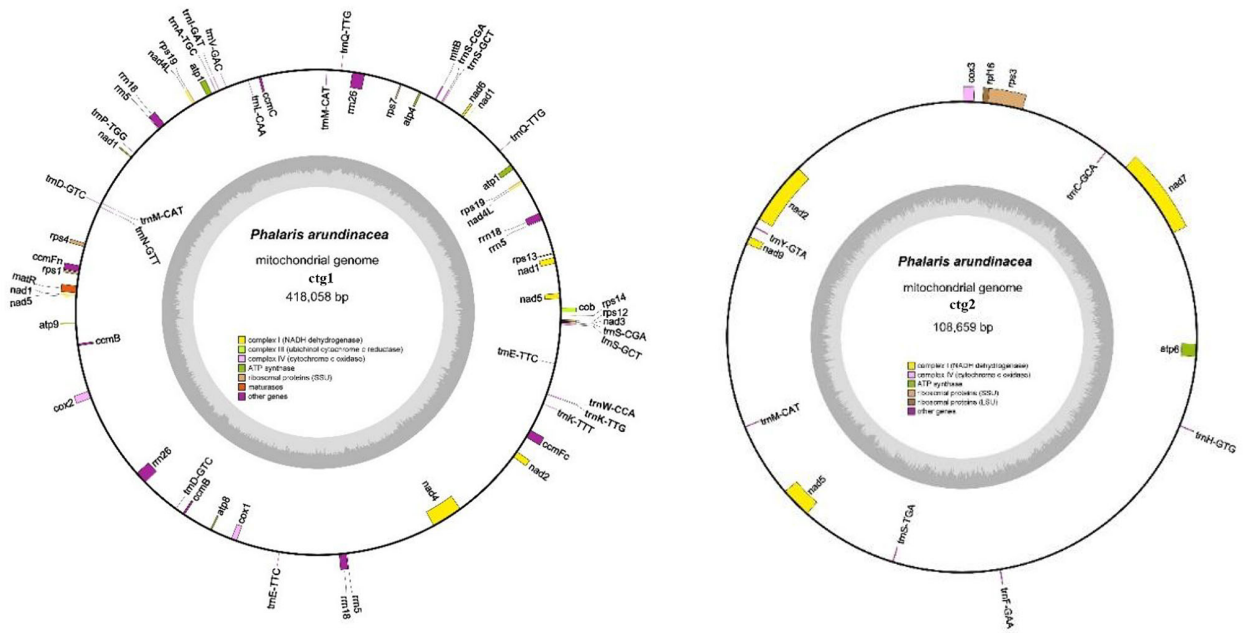


FIGURE 1
Circular representation of the complete mitochondrial genome of *Phalaris arundinacea*. Depicting a circular map, genes on the outer and inner rings are annotated to indicate forward and reverse transcription directions, respectively. The inner ring features a charcoal grey gradient representing GC content, while various functional gene groups are highlighted with distinct colors.

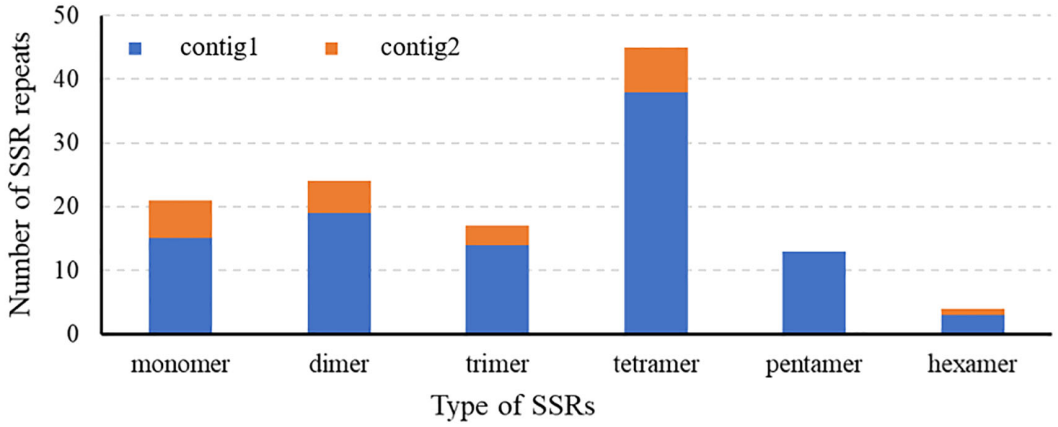


FIGURE 2
Distribution of SSRs in the *P. arundinacea* mitogenome. The X-axis denotes SSR types, and the Y-axis shows the count of repeat units.

trnS-CGA, and *trnS-GCT*). Among PCGs, duplications were restricted to *atp1* and *ccmB*, each present in two copies.

Characterization of repeat sequences

Repeated sequences are widely distributed in the *P. arundinacea* mitogenome and exist as three types: dispersed repeats, tandem repeats, and SSRs (Figure 3). Dispersed repetitive sequences can be classified into four types: forward, palindromic, reverse, and complementary. A total of 318 repetitive sequences (≥28 bp in length) were identified. (Figure 4; Supplementary Table 4). Among these repeats, 154 pairs (48.40%) were forward repeats, with the remainder (164 pairs) being palindromic repeats. Nineteen dispersed repeat sequences, each exceeding 300 bp in length, were identified. The longest forward repeat sequence located in four genes (two *rrn5* and two *rrn18*) measured 19,468 bp, and the longest palindromic repeat, located in ten genes (two *rrn18*, two *atp1*, two *rrn5*, two *nad4L*, and two *rps19*), measured 27,793 bp. In the shorter intervals (12–39 bp and 40–49 bp), there were relatively large numbers of forward and palindromic repeat sequences. As the length increased, the number decreased. However, in some intervals (such as 70–79 bp and 100–199 bp), the number increased.

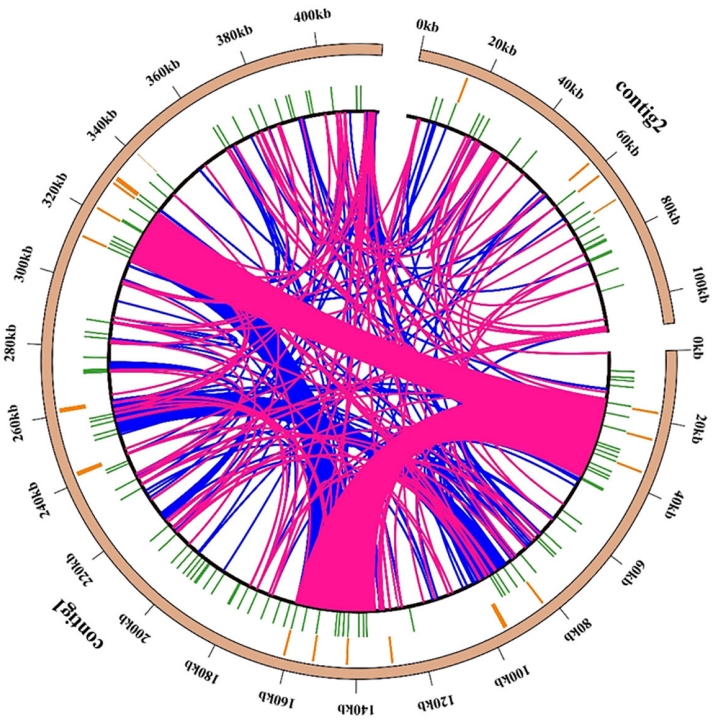


FIGURE 3
Distribution of dispersed repeats in the *Phalaris arundinacea* mitogenome. From the outermost to the innermost, the concentric circles represent the complete mitochondrial genome sequence, tandem repeats, SSR repeats, and dispersed repeats. Dispersed repeats are illustrated by blue arcs (154 forward repeats) and deep pink arcs (164 palindromic repeats).

10.3389/fpls.2026.1926603

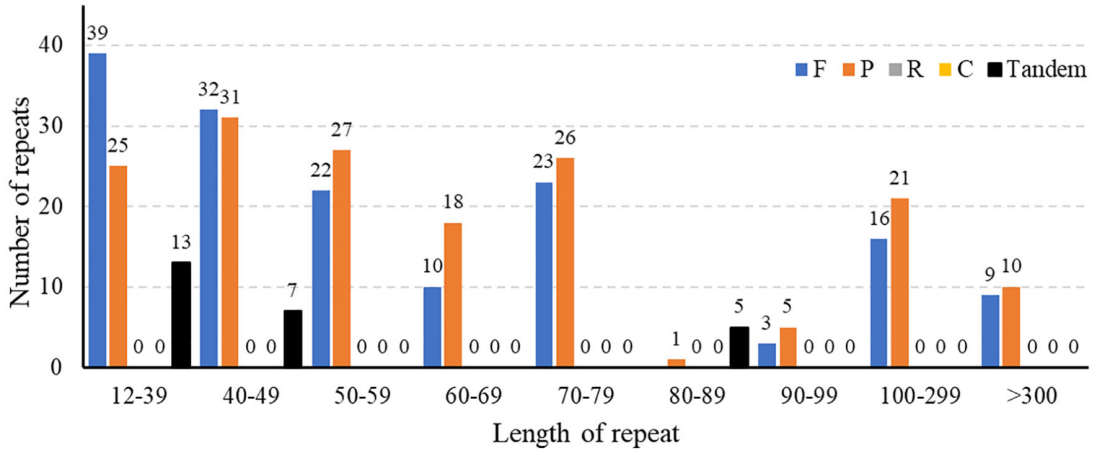


FIGURE 4
Length distribution of dispersed repeats in the *Phalaris arundinacea* mitogenome. The X-axis categorizes the types of dispersed repeats (F, forward; P, palindromic; R, reverse; C, complement), and the Y-axis shows their corresponding counts.

Tandem repeats are present in both prokaryotic and eukaryotic genomes. In the *P. arundinacea* mitogenome, we identified 25 tandem repeats ranging from 24 to 153 bp in length. All these tandem repeats resided in intergenic spacer (IGS) regions, with 52% showing perfect sequence identity (100% match). Their copy number variation ranged from 1.9 to 3.2. (Figure 4; Supplementary Table 8).

A total of 124 SSRs were identified in the *P. arundinacea* mitogenome, distributed on both contigs, of which 102 (82.26%) were found on ctg1 (Supplementary Table 5). Tetramers accounted for the largest proportion of 1–6 nucleotide repeat types, representing 36.29%, of which 38 tetramers were present on ctg1 and seven tetramers were present on ctg2. SSRs in hexameric form were the least distributed on the mitochondrial contigs, accounting for 3.23%, of which three hexamers were present on ctg1 and only one hexamer was present on ctg2. In addition, the pentameric form of the SSRs was found only on ctg1 (Figure 2). We identified 38 different repeating motifs among all SSRs in the *P. arundinacea* mitogenome, including mononucleotides (n = 10.5), dinucleotides (n = 8), trinucleotides (n = 3.4), tetranucleotides (n = 3), pentanucleotides (n = 1.3), and hexanucleotides (n = 1.3). A/T (n = 19) was the most common motif (15.32%), followed by AATG/ATTC (n = 16), 12.90%). While differing in repeat type and motif composition, these sequences shared a strikingly similar genomic arrangement. A total of 116 SSRs (93.55%) were located in the IGS, and a few others were distributed in intron and gene regions (Supplementary Tables 5, S6).

Codon usage analysis of PCGs

RSCU is a metric of evolutionary pressures-including selection, mutation, and drift-acting on codon preference. In this study, we quantified the RSCU values for the *P. arundinacea* mitochondrial genome to assess its codon preference, where a value exceeding 1 denotes a positive bias toward a specific synonymous codon. In total, 37 unique PCGs in the *P. arundinacea* mitogenome are encoded by 9,779 codons. (Supplementary Table 9). The usage frequencies of the amino acids Leu, Ser, and Ile were the highest at 10.84%, 8.89%, and 7.90%, respectively. The frequency of stop codons was the lowest, at only 0.34%. Codon usage analysis confirmed the employment of all 64 codons in the mitogenome, split nearly equally between those with a preferential usage bias (RSCU > 1) and those, conversely, used at below-average frequencies (Figure 5). For example, one alanine codon is encoded by four codons. Codon preference levels varied in the mitogenome of *P. arundinacea*. Specifically, codon GCT exhibited a high usage preference, with an RSCU value of 1.63.

RNA editing sites prediction

RNA editing is a widespread and frequent post-transcriptional process in plant mitochondria. We predicted and analyzed the RNA editing sites of shared PCGs in the mitochondrial genome of *P. arundinacea* and identified 376 valid RNA editing sites located in 31 PCGs. Correspondingly, 23 PCGs were distributed on ctg1 and eight PCGs were distributed on ctg2 (Figure 6). Among these, *ccmC* and *nad7* had the most RNA editing sites (located in ctg1 and ctg2, respectively), both with 27 sites, whereas *mttB* and *rps19* had only one RNA editing site each, both located on ctg1. Eleven types of changes were identified among the 376 RNA editing sites, mainly C-to-T changes (n = 333, 88.56%). Additionally, except for G-to-A changes (n = 15, 3.99%) and C-to-A changes (n = 8, 2.13%), the frequencies of the remaining editing checkpoint changes were all less than 4%. The positional statistics of the RNA editing sites revealed that most of these checkpoints occurred at the second position of the triplet codon (n = 200, 53.19%), followed by the first position (n = 124, 32.98%) (Supplementary Table 10). A total of 49 amino acid variants were found during RNA editing, and 41 amino acid variants appeared on ctg1 (Figure 7), including 10 synonymous substitutions and 3 stop codon changes. The most frequent amino acid substitution involved the conversion of proline to leucine (n = 43), followed by the conversion of serine to leucine (n = 36). In ctg1, the main amino acid variants were proline, leucine, serine, and phenylalanine (57.61%) residues. Although 34 amino acid

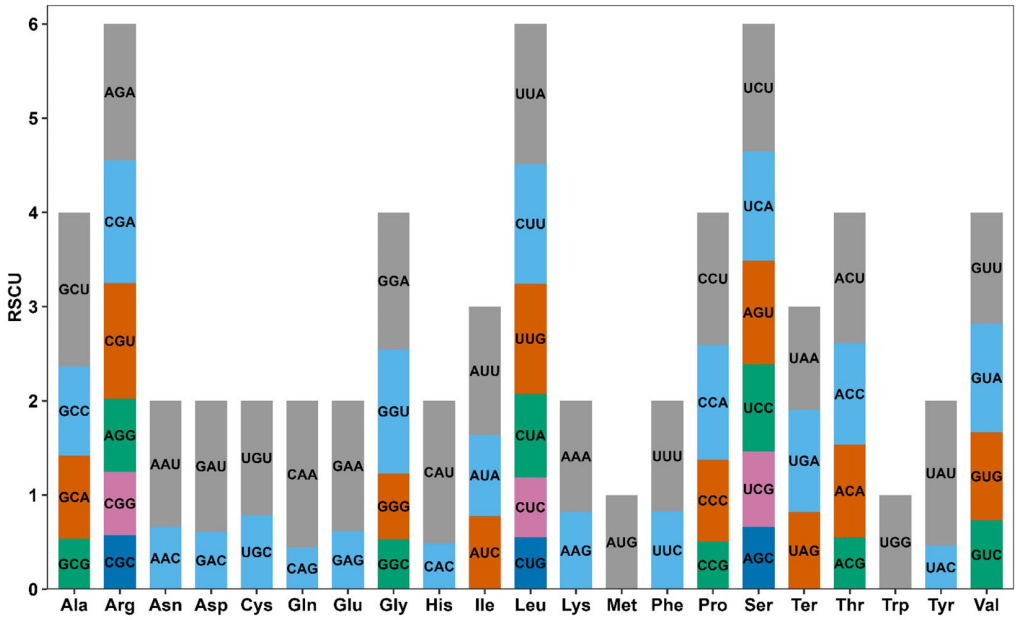


FIGURE 5
RSCU in PCGs of *Phalaris arundinacea* mitogenome. Codon families are on the abscissa.

variants were found on ctg2 (Figure 8), the most frequent substitution observed was from serine to leucine (n = 31); the remaining amino acid variants were all fewer than 14. These results indicate that *P. arundinacea* mitochondrial genomic RNA editing occurs on a relatively small scale and remains generally stable.

Chloroplast-derived mitogenomic sequences

To elucidate the patterns of inter-organellar DNA transfer, homologous regions between the *P. arundinacea* mitochondrial and chloroplast genomes were compared and visualized. (Figure 9). We identified 104 fragments resulting from inter-organellar DNA transfer (Supplementary Table 11). These fragments mainly consist of gene spacers of protein-coding genes and contain small amounts of RNA and tRNA genes. The transferred sequences, spanning 55,866 bp (10.60% of the mitogenome), were overwhelmingly concentrated on ctg1 (86 out of 104 fragments). Notably, the two longest transfer fragments were located on ctg1. One fragment was transferred from the chloroplast (*rrn16, rps12, ndhB (intron), trnL-CAA, rrn23, trnT-CGU, trnA-UGC* (intron), *rps12* (intron), *ndhB, trnV-GAC, rps7, trnT-CGU* (intron), *trnA-UGC*) to the mitochondrial gene spacer (*trnA-UGC, trnL-CAA*). Another fragment was transferred from the chloroplast (*rrn23, trnL-CAA, trnA-UGC* (intron), *trnT-CGU, ndhB* (intron), *rrn16, rps12, trnV-GAC, ndhB, trnA-UGC, rps7, trnT-CGU* (intron), *rps12* (intron)) to the mitochondrial tRNA gene or its codon (*trnA-UGC* (intron), *trnI-GAU, trnL-CAA, trnA-UGC, trnV-GAC, trnI-GAU* (intron)). Both fragments were 10,974 bp in length and exhibited a high degree of sequence similarity (99.927%). In addition, the fragments with the

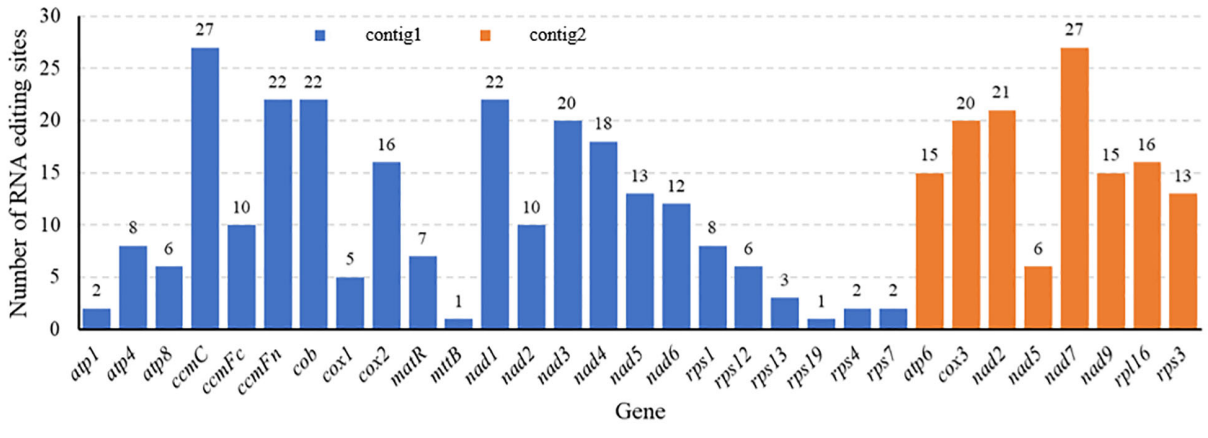


FIGURE 6
RNA editing sites in the *Phalaris arundinacea* mitochondrial PCGs.

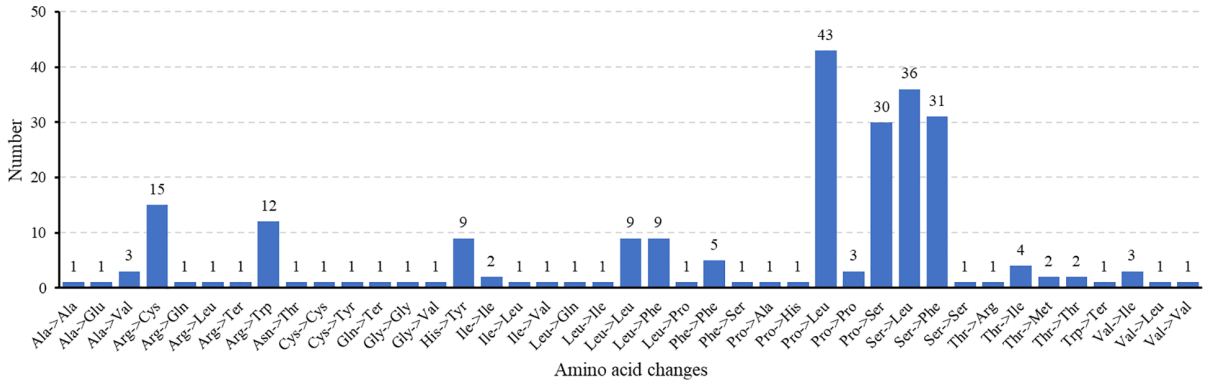


FIGURE 7
RNA editing-induced amino acid substitutions on *P. arundinacea* ctg1.

highest transfer frequency were transferred from *rrn16* and *rrn23* in the chloroplasts to *rrn18* and *rrn26* in the mitochondria (12 and 10 times, respectively).

Selection pressure analysis

Ka/Ks analysis was performed to evaluate the evolutionary selection pressures acting on 27 conserved PCGs across the mitochondrial genomes of 25 Poaceae species and *Phalaris arundinacea* (Figure 10). Most mitochondrial genes exhibited Ka/Ks < 1, suggesting that purifying selection represents the predominant evolutionary force maintaining mitochondrial gene conservation. However, several genes showed lineage-specific elevations in Ka/Ks ratios. In particular, *nad5* displayed Ka/Ks values greater than 1 across the sampled species, while *mttB* and *rps3* exhibited Ka/Ks values greater than 1 in specific lineages, especially within the Poa genus. These elevated Ka/Ks ratios may indicate potential adaptive divergence. Additionally, the Ka/Ks values of *ccmC*, *cox3*, and *rpl16* in *Aegilops cylindrica*, *Aegilops tauschii*, and *Agropyron cristatum* were 0, These values indicate an absence of detected nonsynonymous substitutions in the analyzed alignments, which may reflect strong functional conservation. However, such patterns can also result from limited sequence variation, short aligned regions, or insufficient numbers of informative sites.

Gene loss patterns among the 27 PCGs varied substantially (Supplementary Table 12). *Cynodon dactylon* lacked only *nad4L*, whereas *Eremochloa ophiuroides* and the three *Zea* species lost two genes (*ccmB* and *rpl16*). Extensive gene loss (six genes) occurred in *Aegilops tauschii*, *Elymus magellanicus*, *Fargesia qinlingensis*, and two *Poa* species, likely due to complete sequence conservation or the absence of polymorphic sites. These divergent Ka/Ks dynamics underscore the roles of both purifying and positive selection in shaping mitochondrial genome evolution across Poaceae, thereby providing critical insights into lineage-specific adaptive mechanisms.

Nucleotide diversity

To further elucidate the genetic diversity characteristics of the mitochondrial genome of *Phalaris arundinacea*, we performed nucleotide diversity (Pi) analysis to assess sequence variation across the 30 PCGs (Figure 11; Supplementary Table 13). The Pi values across the PCGs spanned from 0 to 0.38462, with an overall mean of 0.0205, indicating low average nucleotide diversity. Among the PCGs, *rps2* showed the highest Pi value, followed in descending

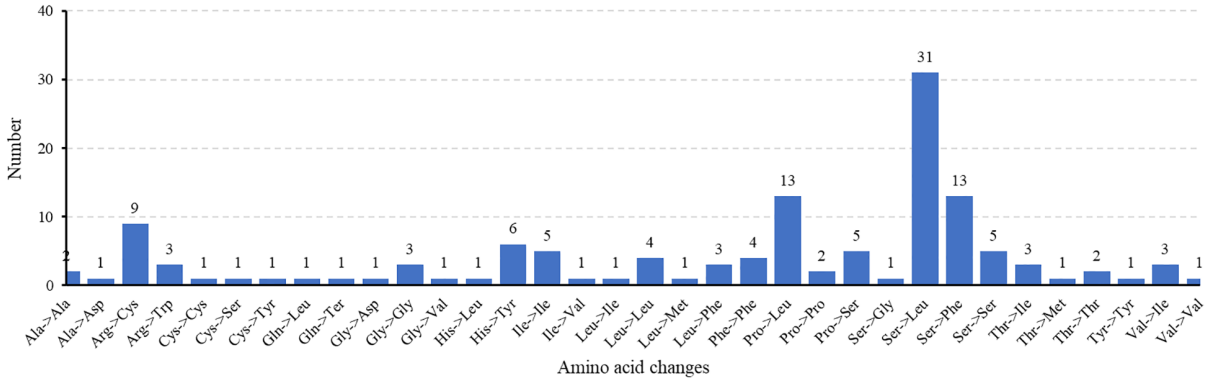


FIGURE 8
RNA editing-induced amino acid substitutions on *P. arundinacea* ctg2.

10.3389/fpls.2026.1926603

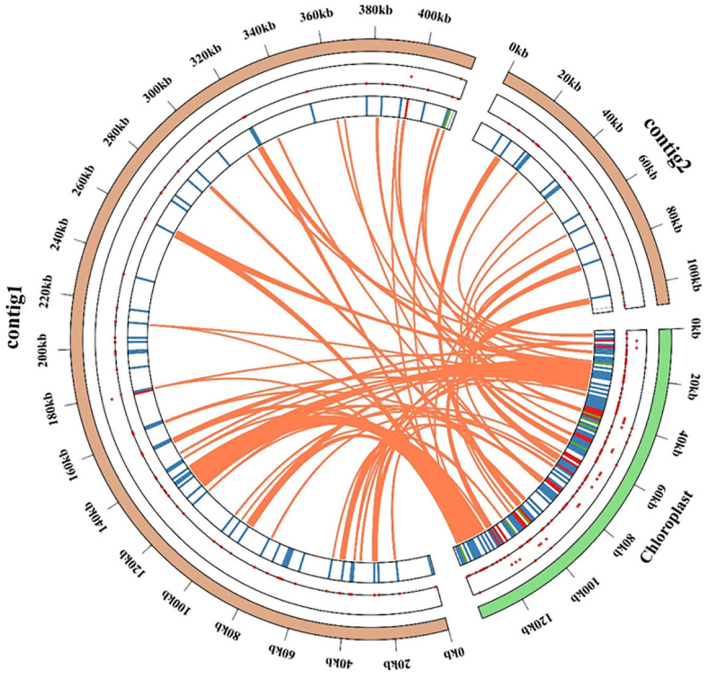


FIGURE 9
Inter-organellar gene transfer between the chloroplast and mitochondrial genomes. The visualization compares the chloroplast genome (light green segment) with the two mitochondrial contigs (purple segments). The internal heatmap density corresponds to the distribution of transferred sequences, with coral lines tracing the insertion routes of chloroplast-derived DNA into the mitogenome.

order by *nad3* and *nad6*, suggesting elevated functional specificity in mitochondrial mRNA translation and NADH dehydrogenase synthesis. In contrast, *ccmFC*, *nad4L*, *rps12*, *rps13*, and *sdh4* displayed Pi values of zero, indicating stringent evolutionary conservation. These conserved genes are critical for driving ATP synthesis and ensuring robust functionality of the electron transport chain.

Phylogenetic analysis

To investigate the evolutionary relationships of *P. arundinacea*, we selected 26 species of Poaceae, including *Phalaris arundinacea*, along with one species of Brassicaceae and one species of Fabaceae. We then integrated published mitochondrial genome data to construct a phylogenetic tree with Brassicaceae and Fabaceae

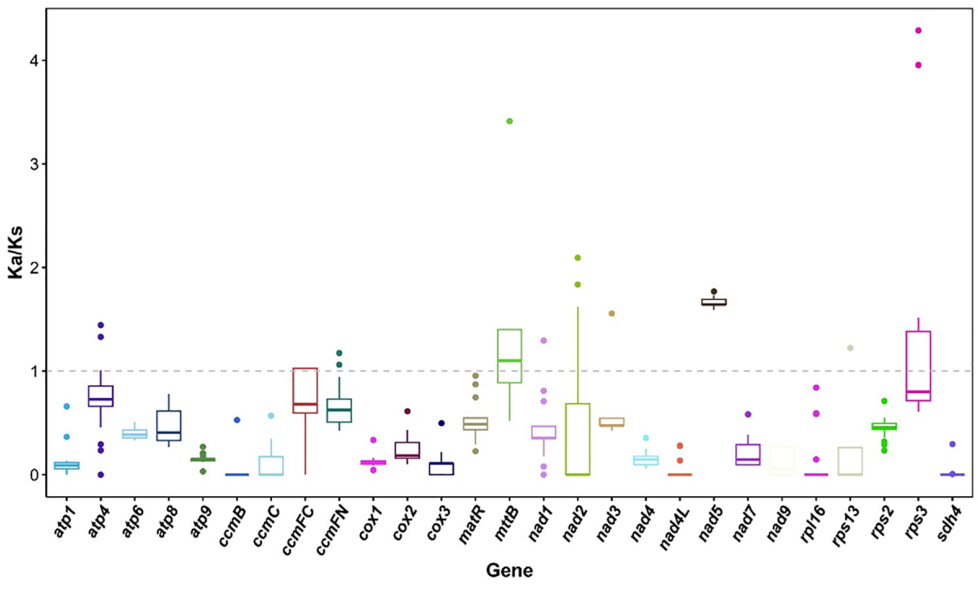


FIGURE 10
Boxplot of Ka/Ks values for 27 mitochondrial PCGs across 25 Poaceae species and *Phalaris arundinacea*.

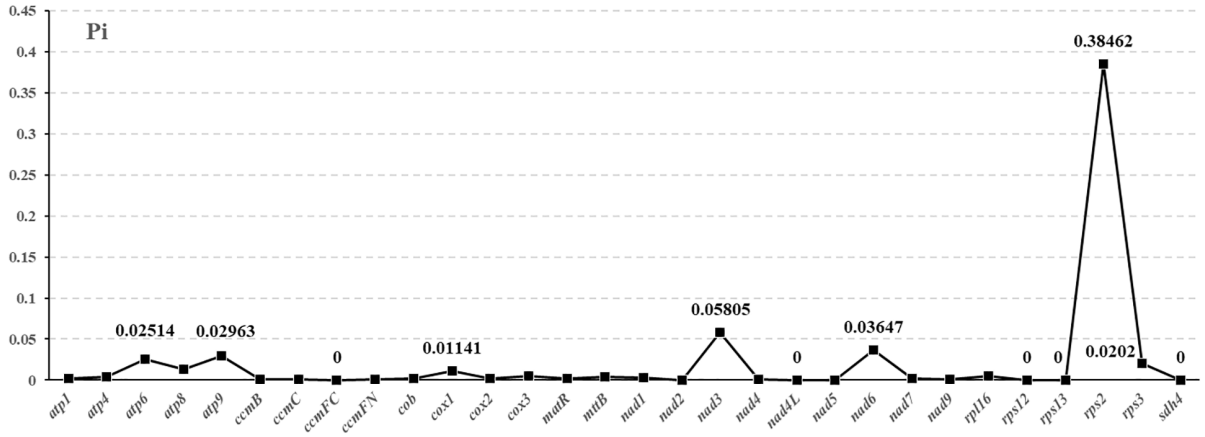


FIGURE 11
Pi of 30 PCGs in the mitochondrial genome of *Phalaris arundinacea*.

species used as outgroups (Figure 12). The analysis revealed that the 26 species could be divided into three groups. The first group comprises *Fargesia, P. arundinacea, Poa, Triticum, Aegilops, Thinopyrum, Agropyron, Elymus, Phragmites, Cynodon, Zea, Chrysopogon, Tripsacum, Sorghum, Eremochloa*, and *Saccharum*. the second group comprises the genus *Oryza*, and the third group comprises the genera *Arabidopsis* and *Medicago*. Among the 25 nodes of the phylogenetic tree, the support rates of 15 nodes exceeded 90%, and 11 reached 100%. This tree clearly distinguished Poaceae plants from the outgroups (Fabaceae and Brassicaceae plants) and differentiated plants of different genera. *P. arundinacea* was identified within the first group. It showed a relatively close relationship with the genus *Poa* and was more closely related to the genera *Triticum, Aegilops, Thinopyrum, Agropyron*, and *Elymus* than to other genera. These findings contribute to resolving Poaceae phylogeny using mitochondrial genomes and provide a framework for investigating evolutionary relationships within the family.

Synteny analysis

Based on the phylogenetic tree results, a comparative synteny analysis was performed between *P. arundinacea* and three closely

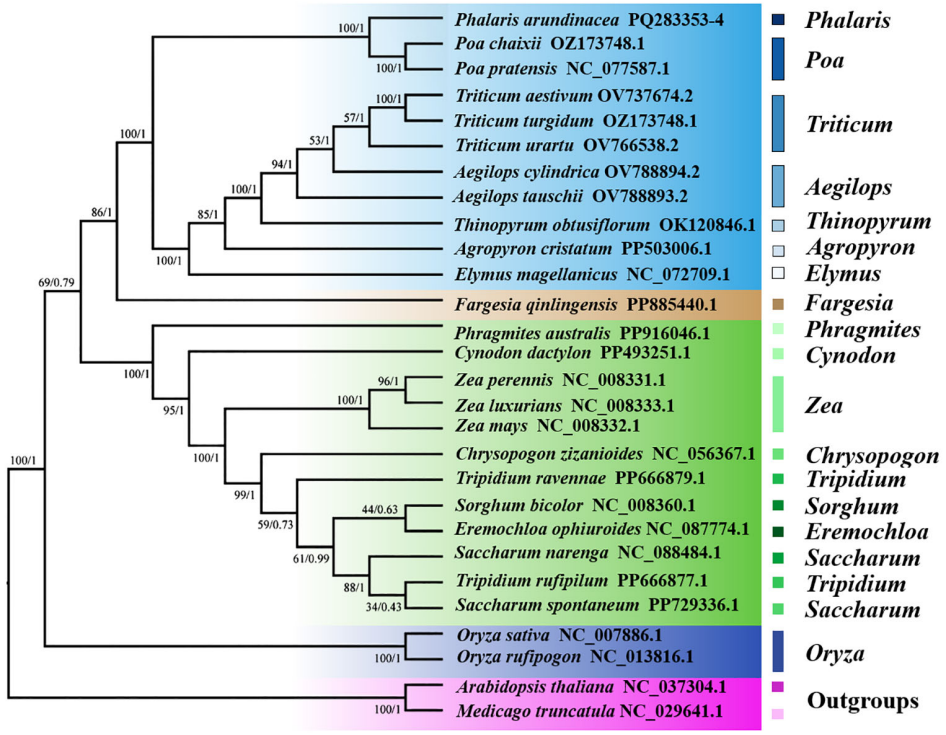


FIGURE 12
Phylogenetic position of *Phalaris arundinacea* among 28 species. Node labels show ML bootstrap values (left) and Bayesian posterior probabilities (right).

related species: *Poa chaixii*, *Poa pratensis*, and *Triticum aestivum* (Figure 13). Pairwise comparisons revealed that 63.84% of syntenic blocks in the mitochondrial genomes of *P. arundinacea* and *Poa chaixii* were homologous. The highest sequence homology (93.33%) was observed between *Poa pratensis* and *Poa chaixii*, thus validating the accuracy of the synteny analysis (Supplementary Table 14).

Discussion

Structural characteristics of the mitogenome in *P. arundinacea*

Resolving the complex evolutionary dynamics of organellar genomes-both plastid and mitochondrial-remains a pivotal challenge in modern phylogenomics. Benefiting from iterative advancements in high-throughput sequencing technologies, extensive high-quality assemblies of nuclear and chloroplast genomes across Poaceae species have been resolved to date. However, the elucidation of Poaceae mitochondrial genomes remains comparatively limited, largely constrained by their complex repeat sequences and frequent structural rearrangements (Gualberto, 2014; Mower, 2020). Most plant research has focused on the chloroplast genome (Liu et al., 2021). Similarly, the chloroplast genome of *Phalaris* has been sequenced (Xiong et al., 2020), but the mitochondrial sequencing of this species has not yet been completed. Accumulating evidence from prior studies has established that plant mitochondrial genomes exhibit an extraordinarily complex genomic architecture. For most land plant taxa, the mitochondrial genome exists *in vivo* in a state of dynamic equilibrium between a full-length master circle and heterogeneous subgenomic circles (Kozik et al., 2019; Tao et al., 2025)- a conformational feature that is postulated to facilitate DNA replication and damage repair or, alternatively, serve as a selective checkpoint during gametogenesis. This study assembled the mitochondrial genome of *P. arundinacea* into two circular structure of 526,717 bp with a GC content of 44.46%, thereby establishing a reference genetic resource for the species. Additionally, the mitochondrial genome of barley (GenBank accession numbers PQ283353 and PQ283354) was 525,599 bp in total length and 44.2% GC (Hisano, 2016), which is very close to that of *Phalaris arundinacea*. Comparative analysis revealed that the *P. arundinacea* mitogenome is larger than that of many sequenced Poaceae species (Supplementary Table 1). This structural feature indicates fewer introns and limited sequence transfers in the *P. arundinacea* mitogenome, which may reflect a more conserved evolutionary trajectory within the Poaceae family. The mitogenome of *P. arundinacea* exhibits a double-ring architecture, as also observed in *Setaria italica* (Zhang, 2024) and *Prunella vulgaris* (Sun et al., 2023). During mitochondrial assembly, the appearance of double-ring structures is caused by the active recombination of repeating sequences (Zhang, 2024). The mitochondrial genome of Ice Grass was reported to contain 35 protein-coding genes (Ou et al., 2024), while 36 were identified in that of *Ventilago leiocarpa* (Guo et al., 2024). By contrast, our annotation of the *P. arundinacea* mitogenome identified 16 core and 21 variable protein-coding genes. The observed disparity could be attributed to evolutionary DNA transfer events among the plastid, nuclear, and mitochondrial compartments. Mitochondria are the dominant source of intracellular reactive oxygen species (ROS), a major driver of DNA damage and mutagenesis, making the mitochondrial genome particularly vulnerable to spontaneous and recurrent mutations (Pakendorf and Stoneking, 2005). The endosymbiotic transfer of core mitochondrial genes to the nuclear genome-which benefits from far more stringent genome surveillance, superior protection from mutagenic insults, and more sophisticated DNA damage repair pathways-is widely

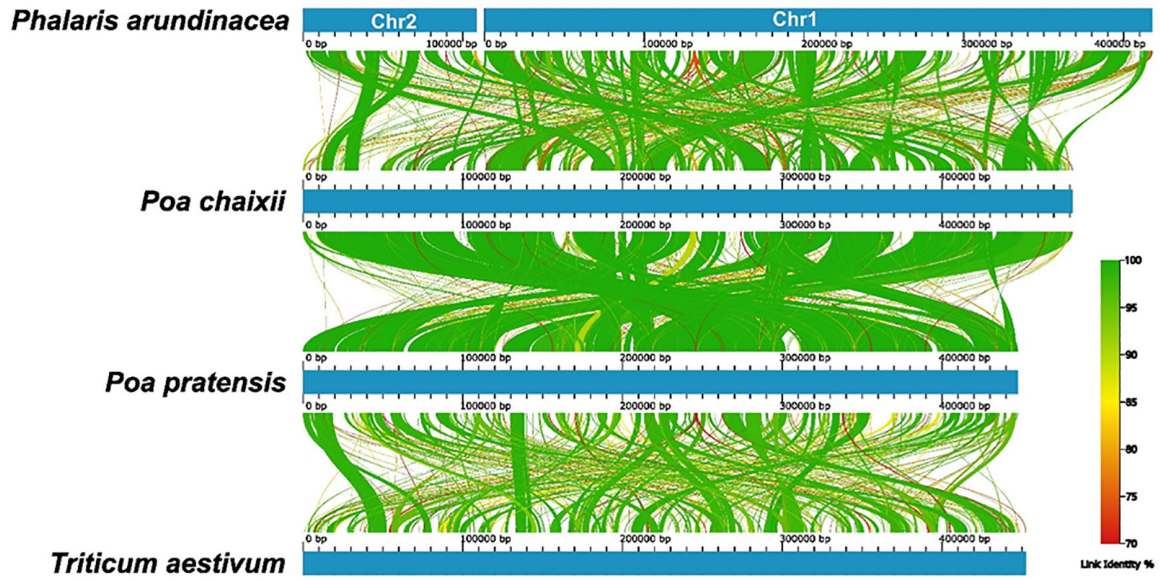


FIGURE 13
Comparative synteny analysis among *Phalaris arundinacea*, *Poa chaixii*, *Poa pratensis*, and *Triticum aestivum*. Sequence identity of syntenic blocks indicated by a green-to-red gradient (high to low).

recognized as a key adaptive evolutionary benefit in eukaryotes (Allen, 2015). Notably, the mitochondrial genome of *P. arundinacea* harbors a reduced complement of annotated genes relative to some other Poaceae lineages. This observation raises the possibility that *P. arundinacea* may have experienced a relatively extensive process of mitochondrial-to-nuclear gene transfer during the course of Poaceae diversification, although alternative explanations - such as gene loss or annotation limitations - cannot be ruled out.

Diversity and variation in the mitogenome of *P. arundinacea*

Our structural dissection of the *P. arundinacea* mitogenome unveiled a substantial proliferation of repetitive elements, spanning simple, tandem, and dispersed iterations. These repetitive sequences are known to drive mitochondrial genome recombination across various scales and at differential abundance (Alverson et al., 2011). Moreover, these repeats can serve as effective molecular markers for genetic analysis owing to their specific sequence signatures (Bi et al., 2016; Ma, 2022). For example, SSRs are often used as molecular markers because of their unique genetic characteristics and lack of environmental influence (Song et al., 2015). In recent years, these markers have been widely used to analyze and identify herbs (Chen, 2022). Mitochondrial genome size shows a significant positive correlation with the count of repeat sequences exceeding 500 bp in length (Allen et al., 2007; Alverson et al., 2024). Through an in-depth investigation of the mitochondrial genome of *Phalaris arundinacea*, we accurately identified 318 repeats. With a total length of 526,717 bp, the mitogenome contained repeat sequences accounting for over 20% of its sequence, highlighting their substantial contribution to its architecture. Illustrating this correlation, the notably larger mitochondrial genome of zucchini (982,833 bp) contains a high density of repeats, accounting for up to 38% of its sequence (Allen et al., 2007). By analyzing the repeats in the mitochondrial genomes of these different species, we further confirmed the close relationship between repeat sequence size and mitochondrial genome size, providing a solid theoretical basis and data support for subsequent research in related fields. The SSR markers identified in this study not only enrich the fundamental genetic dataset for *Phalaris arundinacea* but also provide a robust and directly applicable molecular toolset for three key research and breeding applications: evaluating genetic diversity across wild *P. arundinacea* populations, reconstructing its invasion pathways, and enabling marker-assisted selection in the breeding of elite forage cultivars.

RNA editing constitutes a hallmark feature of plant mitochondrial genomes and can be regarded as template-based post-transcriptional regulation involving single-base switching mode (Zhang, 2024). It can also promote changes in the mitochondrial gene expression patterns (Handa, 2003). A total of 376 RNA editing sites were identified in the *P. arundinacea* mitogenome, a number comparatively fewer than those reported for *Setaria italica* ($n = 417$) (Zhang, 2024) and *Poa pratensis* ($n = 457$) (Ai et al., 2024). In plants, mitochondrial RNA editing depends on the extensively expanded pentatricopeptide repeat (PPR) protein family, whose members bind and modify these target sites with single-nucleotide precision. The primary evolutionarily conserved role of this process is to correct deleterious mutations in the mitochondrial genome, thereby restoring functionally critical conserved amino acid residues in the encoded proteins and maintaining the integrity of mitochondrial respiratory chain function (Steinhauser et al., 1999). The predominant RNA editing event involved C-to-U conversions, primarily at the first and second codon positions. This pattern aligns with the canonical profile reported for plant mitochondrial genomes (Sloan et al., 2010; Small et al., 2020; Ou et al., 2024). RNA editing mainly affected functional genes related to energy metabolism, cytochrome c synthesis, mitochondrial gene expression regulation, and ribosome composition in the mitochondrial genome of *Phalaris arundinacea*, involving a total of 29 protein-coding genes. Relative to the mitochondrial genome of *Agropyron cristatum* (Ou et al., 2024), we detected RNA editing events across a larger set of functional PCGs in the *P. arundinacea* mitogenome. This discrepancy may be attributed to the adaptive evolution of *P. arundinacea* to its native habitats: floodplains, marshes and other wetland environments that are frequently subject to acute hypoxia and submergence stress. The highly efficient RNA editing system serves as a critical post-transcriptional regulatory mechanism that maintains the correct coding sequence of respiratory chain-related genes, thereby ensuring the functional assembly and integrity of the mitochondrial respiratory chain under adverse environmental conditions.

Contrasting with animal mitochondrial genomes, plant mitochondrial genomes display a reduced mutation rate (Darracq et al., 2011) and also incorporate DNA fragments from other cellular compartments, such as the chloroplast (Law et al., 2022). Chloroplast gene fragments have been found in the mitochondrial genomes of many species; however, their transfer varies widely among species (Cheng et al., 2021). In bluegrass, 31 chloroplast-derived segments were transferred to the mitochondrial genome, with 14 of these being complete fragments (Ai et al., 2024); Similarly, in *Suaeda salsa*, 32 mitochondrial genome segments were identified as having originated from the chloroplast genome, encompassing eight annotated genes (Cheng et al., 2021). A substantial amount of chloroplast DNA was identified in the mitogenome, comprising 104 fragments with a cumulative length of 55,866 bp, which accounts for about 10% of the mitochondrial genome sequence, indicating frequent fragment transfer between mitochondria and chloroplasts. However, gene transfers can have different outcomes. Occasionally, the function of the transferred gene remains intact. In other instances, gene structure undergoes modifications. Moreover, there are various types of gene transfer, such as those occurring from entire genes to gene spacer regions.

Evolutionary classification of *P. arundinacea* based on the mitogenome

Codon degeneracy allows multiple nucleotide sequences to encode the same amino acid, while the choice among synonymous codons is often biased. Several factors-gene expression, gene length, tRNA pool, and codon position-collectively contribute to shaping the distinct codon usage bias observed in different organisms (Sun

et al., 2023). Codon usage preferences represent trends in base composition and natural selection in the genome, providing a theoretical basis for species-specific evolution. In the mitochondrial genome of *Phalaris arundinacea*, the codon with the highest usage preference in the mitochondrial genome of *P. arundinacea* was GCT, followed by CAA, TAT, CAT, and TTA, similar to *Agropyron cristatum* (Ou et al., 2024) and *Quercus acutissima* (Liu et al., 2022). Interestingly, codon shifts toward G/C bases were more frequent than those toward A/T. This bias toward higher GC content could contribute to enhanced gene stability. The biological significance of this strong GC preference is not restricted solely to enhancing gene stability. In plant mitochondria, this characteristic is predominantly driven by GC-biased gene conversion (gBGC) (Liu et al., 2020). This is a molecular mechanism tightly coupled to the homologous recombination repair of DNA double-strand breaks (DSBs): when ROS-induced mitochondrial DNA DSBs are repaired via homologous recombination, the mismatch repair system exhibits a consistent insertion preference for G/C bases (Lesecque et al., 2013). The long-term evolutionary accumulation of this effect ultimately shapes the high GC content of the mitochondrial genome and the GC bias in codon usage.

To delineate the selective pressures acting upon the mitochondrial coding sequences, Ka/Ks metrics were calculated, serving as robust indicators of evolutionary constraint versus adaptive divergence. By quantifying the divergence between Ka and Ks substitution rates, this approach enables systematic evaluation of sequence conservation and selection pressure patterns across homologous genes in divergent species (Wang et al., 2010). Evolutionary selection pressure analyses of 27 conserved genes in Poaceae (including *Phalaris arundinacea* and 25 related species) revealed that over 70% of the genes were under strong purifying selection (Ka/Ks < 1), whereas the *nad5* gene exhibited positive selection signals (Ka/Ks > 1) across all species. The *nad5* gene encodes the core transmembrane subunit of mitochondrial respiratory chain Complex I (NADH dehydrogenase), which is responsible for proton translocation across the inner mitochondrial membrane (Tsai et al., 2022). The vast majority of plant respiratory chain genes are under intense purifying selection (Ka/Ks < 1) to maintain essential cellular functions required for organismal survival (Gualberto and Newton, 2017). In *P. arundinacea*, however, *nad5* exhibited Ka/Ks values greater than 1 across the analyzed species, suggesting a possible signal of adaptive divergence. However, elevated Ka/Ks ratios should be interpreted with caution, as they may also arise from stochastic variation, low synonymous substitution rates, alignment uncertainty, or differences in taxonomic sampling. Therefore, although *nad5* represents a potential candidate gene associated with lineage-specific mitochondrial evolution, additional population-level analyses and functional investigations are required to confirm whether adaptive selection contributed to its diversification. The functional significance of the observed nad5 variation remains unclear and requires further investigation using broader taxonomic sampling and functional validation. This finding provides a key target for future research on energy metabolism in *Poaceae* plants. Furthermore, synonymous substitutions predominate in plant mitochondrial genomes, and genetic variations in critical functional genes are strictly constrained by purifying selection, indicating their functional irreplaceability (Ou et al., 2024). Ka/Ks ratio analysis enabled the precise identification of adaptive divergence patterns in orthologous PCGs across species, providing molecular evidence for deciphering lineage-specific evolutionary trajectories.

To validate the taxonomy of *Phalaris* plants using their mitochondrial genome, we constructed a phylogenetic tree containing 28 plant mitochondrial genomes. Earlier studies classified *Phalaris* within the oat-bluegrass complex of the subfamily Pooideae (Quintanar et al., 2007; Schneider et al., 2009; Saarela et al., 2015), In recent years, *Phalaris arundinacea* has been reported to possess multiple ploidy levels, although tetraploid dominates. As a result, some taxonomists have recognized two taxonomic categories of *Phalaris arundinacea*: subspecies and variants (Perdereau et al., 2017). Some scholars believe that *Ph*alaris exhibits aneuploidy, some forms of which may have evolved through diploidy. The genus lacks subordinate classification, and species relationships and subordinate divisions should be evaluated using with multidisciplinary information, such as morphology (Handa, 2003; Voshell and Hilu, 2014)Phylogenetic analyses revealed that *Phalaris* is relatively closely related to *Poa* and forms a separate branch, which is consistent with previous results. Furthermore, comparative synteny analysis provided structural genomic evidence supporting this phylogenetic relationship. The results revealed that 63.84% of the syntenic blocks between the mitochondrial genomes of *P. arundinacea* and *Poa chaixii* were homologous. This finding confirms a hallmark evolutionary characteristic of plant mitochondrial genomes: an extremely low nucleotide substitution rate coupled with an exceptionally high structural rearrangement rate (Palmer and Herbon, 1988). Frequent structural rearrangements also act as the underlying cytological driver for the occurrence of polyploidy and the complex infrageneric classification within the *Phalaris* genus. The highest sequence homology (93.33%) was observed between *Poa pratensis* and *Poa chaixii* within the genus *Poa*, validating the accuracy of the current synteny analysis. These findings support the close phylogenetic relationship between *P. arundinacea* and species of the genus Poa. The mitochondrial genome data generated in this study provide a valuable reference and a foundation for future phylogenetic studies within Phalaris that incorporate additional chloroplast, nuclear, or mitochondrial genome data from a broader representation of Phalaris species.

Conclusions

This study presents the first assembly and structural characterization of the *P. arundinacea* mitochondrial genome, a circular molecule of 526,717 bp harboring 72 genes (37 PCGs, 27 tRNAs, and 8 rRNAs). Subsequently, we investigated the dynamic aspects of the *P. arundinacea* mitogenome by analyzing repeat sequences, RNA editing sites, and inter-organellar gene transfer. Furthermore, through analyses of codon usage bias, nucleotide diversity, and phylogenetics, we assessed its evolutionary trajectory and phylogenetic placement within the Poaceae. Collectively, this work provides foundational insights into the genetic evolution of *P. arundinacea*

and addresses a key gap in its genomic resources. It also provides a solid theoretical foundation for subsequent genetic breeding exploration and offers valuable scientific guidelines for the mining and application of the valuable genetic characteristics contained in *P. arundinacea*.

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

ZT: Writing – original draft, Data curation, Writing – review & editing. CT: Writing – review & editing, Resources, Investigation. YY: Investigation, Writing – review & editing. WG: Methodology, Formal analysis, Writing – review & editing. ZL: Writing – review & editing, Supervision, Project administration. XB: Visualization, Writing – review & editing. ZW: Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. The research was funded by Central Public-interest Scientific Institution Basal Research Fund (No. Y2026YC67), Science & Technology Fundamental Resources Investigation Program (2024FY100302), Agricultural Science and Technology Innovation Program of CAAS (No. 27-GRI-01), the Central Public-interest Scientific Institution Basal Research Fund (1610332023003), the Natural Science Foundation of Inner Mongolia Autonomous Region (2025QN03151), the Inner Mongolia Science and Technology Plan (2025YFDZ0063).

Conflict of interest

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

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2026.1926603/full#supplementary-material

SUPPLEMENTARY TABLE 1
GenBank accession numbers of plant mitogenomes sampled in this study.

SUPPLEMENTARY TABLE 2
Gene composition in the *Phalaris arundinacea* mitogenome.

SUPPLEMENTARY TABLE 3
The start codon and stop codon information of *Phalaris arundinacea* mitochondrial gene.

SUPPLEMENTARY TABLE 4
Size of introns and exons in split genes in *Phalaris arundinacea* mitogenome.

SUPPLEMENTARY TABLE 5
Microsatellite repeats in *Phalaris arundinacea* mitogenome.

SUPPLEMENTARY TABLE 6
Distribution of SSRs in *Phalaris arun*dinacea mitogenome.

SUPPLEMENTARY TABLE 7
Dispersed repeats in *Phalaris arundinacea* mitogenome.

SUPPLEMENTARY TABLE 8
Distribution of tandem repeats in *Phalaris arundinacea* mitogenome.

SUPPLEMENTARY TABLE 9
Codon counts in the Phalaris arundinacea mitochondrial PCGs.

SUPPLEMENTARY TABLE 10
RNA editing sites in the mitogenome of *Phalaris arundinacea*.

SUPPLEMENTARY TABLE 11
Fragments transferred from chloroplast to mitochondria in *Phalaris arundinacea*.

SUPPLEMENTARY TABLE 12
The pairwise Ka/Ks values among the 25 Poaceae species.

SUPPLEMENTARY TABLE 13
The genetic diversity characteristics of the mitochondrial genome in *Phalaris arundinacea*.

SUPPLEMENTARY TABLE 14
The collinearity proportion of the genomes of four species.

References

Ai, Y., Chen, J., Chen, Y., Wang, X., Wang, N., Liang, W., et al. (2024). Assembly and comparative analysis of the complete mitochondrial genome of *poa pratensis*. *Crop Sci.* 64, 244–259. doi: 10.1002/csc2.21147

Allen, J. F. (2015). Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. *Proc. Natl. Acad. Sci.* 112, 10231–10238. doi: 10.1073/pnas.1500012112

Allen, J. O., Fauron, C. M., Minx, P., Roark, L., Oddiraju, S., Lin, G. N., et al. (2007). Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. *Genetics* 177, 1173–1192. doi: 10.1534/genetics.107.073312

Alverson, A. J., Wei, X., Rice, D. W., Stern, D. B., Barry, K., and Palmer, J. D. (2024). Insights into the evolution of mitochondrial genome size from complete sequences of citrullus lanatus and cucurbita pepo (cucurbitaceae). 27(6), 1436–1448. doi: 10.1093/molbev/msq029

Alverson, A. J., Zhuo, S., Rice, D. W., Sloan, D. B., and Palmer, J. D. (2011). The mitochondrial genome of the legume vigna radiata and the analysis of recombination across short mitochondrial repeats. *PloS One* 23, 2499–2513. doi: 10.1371/journal.pone.0016404

Ankenbrand, M. J., Hohlfeld, S., Hackl, T., and Förster, F. (2017). AliTV—interactive visualization of whole genome comparisons. *PeerJ Comput. Sci.* 3, e116. doi: 10.7717/peerj-cs.116

Apfelbaum, S. I., and Sams, C. E. (1987). Ecology and control of reed canary grass (Phalaris arundinacea L.). 7.

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

Benson, G. (1999). Tandem repeats finder: A program to analyze DNA sequences. 27, 12–17. doi: 10.1093/nar/27.2.573

Bergthorsson, U., Adams, K. L., Thomason, B., and Palmer, J. D. (2003). Widespread horizontal transfer of mitochondrial genes in flowering plants 424, 424. doi: 10.1038/nature01743

Bi, C., Paterson, A. H., Wang, X., Xu, Y., Wu, D., Qu, Y., et al. (2016). Analysis of the complete mitochondrial genome sequence of the diploid cotton gossypium raimondii by comparative genomics approaches. *BioMed. Res. Int.* 2016(1). doi: 10.1155/2016/5040598

Chen, Y., Ye, W., Zhang, Y., and Xu, Y. (2015). High speed BLASTN: An accelerated MegaBLAST search tool. *Nucleic Acids Res.* 43, 7762–7768. doi: 10.1093/nar/gkv784

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, S. (2022). Emerging biotechnology applications in natural product and synthetic pharmaceutical analyses. 12(11), 4075–4097. doi: 10.1016/j.apsb.2022.08.025

Chen, X. J., and Butow, R. A. (2005). The organization and inheritance of the mitochondrial genome. *Nat. Rev. Genet.* 6, 815–825. doi: 10.1038/nrg1708

Cheng, Y., He, X., Priyadarshani, S. V. G. N., Wang, Y., Ye, L., Shi, C., et al. (2021). Assembly and comparative analysis of the complete mitochondrial genome of suaeda glauca. *BMC Genomics* 22, 167. doi: 10.1186/s12864-021-07490-9

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* 10, giab008. doi: 10.1093/gigascience/giab008

Darracq, A., Varre, J. S., Saumitou-Laprade, P., Oztas, S., Lenoble, P., Vacherie, B., et al. (2011). Structural and content diversity of mitochondrial genome in beet: A comparative genomic analysis. *Genome Biol. Evol.* 3, 723–736. doi: 10.1093/gbe/evr042

Galatowitsch, S. M., Anderson, N. O., and Ascher, P. D. (1999). Invasiveness in wetland plants in temperate north America 19, 19. doi: 10.1007/bf03161781

Greiner, S., Lehwark, P., and Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47, W59–W64. doi: 10.1093/nar/gkz238

Gualberto, J. M. (2014). The plant mitochondrial genome: Dynamics and maintenance. 100, 107–120. doi: 10.1016/j.biochi.2013.09.016

Gualberto, J. M., and Newton, K. J. (2017). Plant mitochondrial genomes: dynamics and mechanisms of mutation. *Annu. Rev. Plant Biol.* 68, 225–252. doi: 10.1146/annurev-arplant-043015-112232

Guo, S., Li, Z., Li, C., Liu, Y., Liang, X., and Qin, Y. (2024). Assembly and characterization of the complete mitochondrial genome of ventilago leiocarpa. *Plant Cell Rep.* 43, 77. doi: 10.1007/s00299-023-03126-2

Handa, H. (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (brassica napus L.): Comparative analysis of the mitochondrial genomes of rapeseed and arabidopsis thaliana. *Nucleic Acids Res.* 31, 5907–5916. doi: 10.1093/nar/gkg795

Hisano, H. (2016). Mitochondrial genome sequences from wild and cultivated barley (hordeum vulgare). 17, 824. doi: 10.1186/s12864-016-3159-3

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754

Iriarte, A., Lamolle, G., and Musto, H. (2021). Codon usage bias: An endless tale. *J. Mol. Evol.* 89, 589–593. doi: 10.1007/s00239-021-10027-z

Jin, J.-J. (2020). GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. 21, 241. doi: 10.1186/s13059-020-02154-5

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

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive *k* -mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116

Kozik, A., Rowan, B. A., Lavelle, D., Berke, L., Schranz, M. E., Michelmore, R. W., et al. (2019). The alternative reality of plant mitochondrial DNA: one ring does not rule them all. *PloS Genet.* 15, e1008373. doi: 10.1371/journal.pgen.1008373

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. 29, 4633–4642. doi: 10.1093/nar/29.22.4633

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. 9, 357–359. doi: 10.1038/nmeth.1923

Lavergne, S., Muenke, N. J., and Molofsky, J. (2010). Genome size reduction can trigger rapid phenotypic evolution in invasive plants. *Annals of Botany*, 105(1), 109–116. doi: 10.1093/aob/mcp271

Lavergne, S., and Molofsky, J. (2004). Reed canary grass (phalaris arundinacea) as a biological model in the study of plant invasions. 23(5), 415–429. doi: 10.1080/07352680490505934

Law, S. S. Y., Liou, G., Nagai, Y., Giménez-Dejoz, J., Tateishi, A., Tsuchiya, K., et al. (2022). Polymer-coated carbon nanotube hybrids with functional peptides for gene delivery into plant mitochondria. *Nat. Commun.* 13, 2417. doi: 10.1038/s41467-022-30185-y

Lesecque, Y., Mouchiroud, D., and Duret, L. (2013). GC-biased gene conversion in yeast is specifically associated with crossovers: molecular mechanisms and evolutionary significance. *Mol. Biol. Evol.* 30, 1409–1419. doi: 10.1093/molbev/mst056

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with burrows–wheeler transform. 26, 589–595. doi: 10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, J., Ni, Y., Lu, Q., Chen, H., and Liu, C. (2025). PMGA: A plant mitochondrial genome annotator. *Plant Commun.* 6, 101191. doi: 10.1016/j.xplc.2024.101191

Liu, F., Fan, W., Yang, J.-B., Xiang, C.-L., Mower, J. P., Li, D.-Z., et al. (2020). Episodic and guanine–cytosine-biased bursts of intragenomic and interspecific synonymous divergence in ajugoideae (lamiaceae) mitogenomes. *New Phytol.* 228, 1107–1114. doi: 10.1111/nph.16753

Liu, D., Guo, H., Zhu, J., Qu, K., Chen, Y., Guo, Y., et al. (2022). Complex physical structure of complete mitochondrial genome of quercus acutissima (fagaceae): A significant energy plant. 13(8), 1321. doi: 10.3390/genes13081321

Liu, K., Wang, R., Guo, X.-X., Zhang, X.-J., Qu, X.-J., and Fan, S.-J. (2021). Comparative and phylogenetic analysis of complete chloroplast genomes in eragrostideae (chloridoideae, poaceae). *Plants* 10, 109. doi: 10.3390/plants10010109

Ma, Q. (2022). Assembly and comparative analysis of the first complete mitochondrial genome of. 22, 29. doi: 10.1186/s12870-021-03416-5

Mower, J. P. (2020). Variation in protein gene and intron content among land plant mitogenomes. 53, 203–213. doi: 10.1016/j.mito.2020.06.002

Ou, T., Wu, Z., Tian, C., Yang, Y., and Li, Z. (2024). Complete mitochondrial genome of agropyron cristatum reveals gene transfer and RNA editing events. 24, 830. doi: 10.1186/s12870-024-05558-8

Pakendorf, B., and Stoneking, M. (2005). Mitochondrial dna and human evolution. *Annu. Rev. Genomics Hum. Genet.* 6, 165–183. doi: 10.1146/annurev.genom.6.080604.162249

Palmer, J. D., and Herbon, L. A. (1988). Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. *J. Mol. Evol.* 28, 87–97. doi: 10.1007/BF02143500

Perdereau, A., Klaas, M., Barth, S., and Hodkinson, T. R. (2017). Plastid genome sequencing reveals biogeographical structure and extensive population genetic variation in wild populations of *phalaris arundinacea* L. in north-western europe. *GCB Bioenergy* 9, 46–56. doi: 10.1111/gcbb.12362

Qu, X.-J., Moore, M. J., Li, D.-Z., and Yi, T.-S. (2019). PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* 15, 50–62. doi: 10.1186/s13007-019-0435-7

Quintanar, A., Castroviejo, S., and Catalán, P. (2007). Phylogeny of the tribe aveneae (pooideae, poaceae) inferred from plastid *trnT-F* and nuclear ITS sequences. *Am. J. Bot.* 94, 1554–1569. doi: 10.3732/ajb.94.9.1554

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, 3299–3302. doi: 10.1093/molbev/msx248

Saarela, J. M., Wysocki, W. P., Barrett, C. F., Soreng, R. J., Davis, J. I., Clark, L. G., et al. (2015). Plastid phylogenomics of the cool-season grass subfamily: clarification of relationships among early-diverging tribes. *AoB Plants* 7, plv046. doi: 10.1093/aobpla/plv046

Sahlin, K. (2023). A survey of mapping algorithms in the long-reads era. 24, 133. doi: 10.1186/s13059-023-02972-3

Sahramaa, M., Hömmö, L., and Jauhiainen, L. (2004). Variation in seed production traits of reed canarygrass germplasm. *Crop Sci.* 44, 988–996. doi: 10.2135/cropsci2004.9880

Schneider, J., Döring, E., Hilu, K. W., and Röser, M. (2009). Phylogenetic structure of the grass subfamily pooideae based on comparison of plastid *matK* gene–3′ *trnK* exon and nuclear ITS sequences. *Taxon* 58, 405–424. doi: 10.1002/tax.582008

Shi, L., Chen, H., Jiang, M., Wang, L., Wu, X., Huang, L., et al. (2019). CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47, W65–W73. doi: 10.1093/nar/gkz345

Sloan, D. B. (2013). One ring to rule them all? Genome sequencing provides new insights into the 'master circle' model of plant mitochondrial *DNA* structure. *New Phytol.* 200, 978–985. doi: 10.1111/nph.12395

Sloan, D. B., MacQueen, A. H., Alverson, A. J., Palmer, J. D., and Taylor, D. R. (2010). Extensive loss of RNA editing sites in rapidly evolving silene mitochondrial genomes: selection vs. retroprocessing as the driving force. 27, 2447–2456. doi: 10.1534/genetics.110.118000

Small, I. D., Schallenberg-Rüdinger, M., Takenaka, M., Mireau, H., and Ostersetzer-Biran, O. (2020). Plant organellar RNA editing: what 30 years of research has revealed. *Plant J.* 101, 1040–1056. doi: 10.1111/tpj.14578

Song, X., Ge, T., Li, Y., and Hou, X. (2015). Genome-wide identification of SSR and SNP markers from the non-heading Chinese cabbage for comparative genomic analyses. 16, 328. doi: 10.1186/s12864-015-1534-0

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446

Steinhauser, S., Beckert, S., Capesius, I., Malek, O., and Knoop, V. (1999). Plant mitochondrial RNA editing. *J. Mol. Evol.* 48, 303–312. doi: 10.1007/PL00006473

Sun, Z., Wu, Y., Fan, P., Guo, D., Zhang, S., and Song, C. (2023). Assembly and analysis of the mitochondrial genome of *Prunella vulgaris*. *Front. Plant Sci.* 14, 1237822. doi: 10.3389/fpls.2023.1237822

Tao, A., Tao, J., Gan, Z., Zhang, Y., Yin, C., and Jiang, G. (2025). The characterisation of the complete mitochondrial genome of polygonatum kingianum reveals recombination mediated by repeats associated with DNA replication. *Sci. Rep.* 15, 19980. doi: 10.1038/s41598-025-01756-y

Thomsen, M., Brownell, K., Groshek, M., and Kirsch, E. (2012). Control of reed canarygrass promotes wetland herb and tree seedling establishment in an upper mississippi river floodplain forest. doi: 10.1007/s13157-012-0289-5

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. doi: 10.1093/nar/gkx391

Tsai, H.-C., Hsieh, C.-H., Hsu, C.-W., Hsu, Y.-H., and Chien, L.-F. (2022). Cloning and organelle expression of bamboo mitochondrial complex I subunits Nad1, Nad2, Nad4, and Nad5 in the yeast saccharomyces cerevisiae. *Int. J. Mol. Sci.* 23, 4054. doi: 10.3390/ijms23074054

Turmel, M., Otis, C., and Lemieux, C. (2016). Mitochondrion-to-chloroplast DNA transfers and intragenomic proliferation of chloroplast group II introns in gloeotilopsis green algae (ulotrichales, ulvophyceae). *Genome Biol. Evol.* doi: 10.1093/gbe/evw190

Voshell, S. M., and Hilu, K. W. (2014). Canary *G* rasses (*P halaris* , P oaceae): biogeography, molecular dating and the role of floret structure in dispersal. *Mol. Ecol.* 23, 212–224. doi: 10.1111/mec.12575

Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinf.* 8, 77–80. doi: 10.1016/S1672-0229(10)60008-3

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, e1005595. doi: 10.1371/journal.pcbi.1005595

Wick, R. R., Schultz, M. B., Zobel, J., and Holt, K. E. (2015). Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31, 3350–3352. doi: 10.1093/bioinformatics/btv383

Wu, S., Liu, W., Aljohi, H. A., Alromaih, S. A., Alanazi, I. O., Lin, Q., et al. (2018). REDO: RNA editing detection in plant organelles based on variant calling results. *J. Comput. Biol.* 25, 509–516. doi: 10.1089/cmb.2017.0214

Wynn, E. L., and Christensen, A. C. (2019). Repeats of unusual size in plant mitochondrial genomes: identification, incidence and evolution. doi: 10.1534/g3.118.200948

Xiong, Y., Xiong, Y., Jia, S., and Ma, X. (2020). The complete chloroplast genome sequencing and comparative analysis of reed canary grass (phalaris arundinacea) and hardinggrass (P. aquatica). doi: 10.3390/plants9060748

Zhang, Z. (2022). KaKs_Calculator 3.0: calculating selective pressure on coding and non-coding sequences. *Genomics Proteomics Bioinf.* 20, 536–540. doi: 10.1016/j.gpb.2021.12.002

Zhang, J. (2024). Assembly and comparative analysis of the first complete mitochondrial genome of setaria italica. doi: 10.1007/s00425-024-04386-2

Zou, Y., Zhu, W., Hou, Y., Sloan, D. B., and Wu, Z. (2025). The evolutionary dynamics of organellar pan-genomes in arabidopsis thaliana. *Genome Biol.* 26, 240. doi: 10.1186/s13059-025-03717-0