



OPEN ACCESS

EDITED BY

Zhiqiang Wu,
Chinese Academy of Agricultural Sciences,
China

REVIEWED BY

Jie Wang,
Murdoch University, Australia
Yoshinori Fukasawa,
Utsunomiya University, Japan
Quan Kuang,
Nanchang Normal University, China
Rongxiang Zhang,
Guizhou Education University, China

*CORRESPONDENCE

Zinian Wu
✉ wuzinian@caas.cn
Zhiyong Li
✉ lizhiyong01@caas.cn

RECEIVED 17 June 2025

ACCEPTED 18 August 2025

PUBLISHED 05 September 2025

CITATION

Tian Y, Wu Z, Tian C, Yang Y and Li Z (2025)
Phylogenetic classification and genetic
insights from the complete mitochondrial
genome of *Medicago ruthenica*.
Front. Plant Sci. 16:1648505.
doi: 10.3389/fpls.2025.1648505

COPYRIGHT

© 2025 Tian, Wu, Tian, Yang and Li. 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.

Phylogenetic classification and genetic insights from the complete mitochondrial genome of *Medicago ruthenica*

Yonglei Tian^{1,2}, Zinian Wu^{1*}, Chunyu Tian^{1,2}, Yanting Yang^{1,2}
and Zhiyong Li^{2*}

¹Institute of Grassland Research, Chinese Academy of Agricultural Sciences, Hohhot, China, ²Key
Laboratory of Grassland Resources and Utilization of Ministry of Agriculture, Hohhot, China

Introduction: *Medicago ruthenica*, an essential wild perennial relative of *Medicago sativa*, is distributed across Eurasia, with good palatability and strong tolerance to drought, cold, and saline-alkali stress. It exhibits wide adaptability, and its morphology varies greatly under different environmental conditions. Owing to this rich morphological diversity, the inference of its phylogeny has been controversial. Extensive studies have been conducted on the nuclear and chloroplast genomes of *M. ruthenica*; however, because of the complexity of the mitochondrial genome, the *M. ruthenica* mitochondrial genome has not been systematically described.

Methods: In this study, we successfully assembled the *M. ruthenica* mitochondrial genome and investigated various related aspects, including genomic features, RNA editing sites, codon preference, gene transfer events, phylogeny, selective pressure, nucleotide diversity analysis and collinearity analysis.

Results and discussion: The mitogenome of *M. ruthenica* is a circular DNA molecule measuring 354,988 base pairs (bp) in length and has a GC content of 45.13%. The mitochondrial genome of *M. ruthenica* includes 55 genes, with 34 being unique protein-coding genes (PCGs) (excluding three likely pseudogenes), along with three ribosomal RNA genes and 18 transfer RNA genes. The mitogenome is abundant with repetitive sequences; a total of 148 dispersed repeats were identified, including 79 palindromic repeats (53%) and 69 forward repeats (47%), without inverted or complementary repeats. Leucine (Leu) was the most frequently occurring amino acid in PCGs at 10.63%, while tryptophan (Trp) was the least frequent at 1.46%. The codon AUU (Ile) had the highest frequency of use, with the relative synonymous codon usage (RSCU) value of 1.29. *atp8*, *atp4*, *rps4*, *nad4*, *ccmB*, *ccmFc* and *nad9* showed significant positive selection pressure in multiple legumes (Ka/Ks > 1). Additionally, *M. ruthenica*'s phylogenetic relationship was categorized under *Medicago*, supported by mitogenome data from 33 other legume species. Understanding the mitochondrial genome characteristics of *M. ruthenica* is critical for elucidating its genetic background. The results of this study serve as a basis for the genetic breeding of *M. ruthenica*.

KEYWORDS

Medicago ruthenica, mitochondrial genome, gene transfer, RNA editing, phylogenetic analysis

1 Introduction

Medicago ruthenica is a perennial forage legume distributed across Eurasia (Small and Jomphe, 1989). Widely recognized as a premium protein forage, it is valued for its resilience to cold, drought and barren soils, as well as its trampling tolerance. Characterized by large leaves and soft stems, this forage has exceptional nutritional value and palatability. Notably, it contains no saponins and poses no risk of bloating disease in livestock, making it a highly desirable feed source (Choi et al., 2020; Shu et al., 2018). The crown of *M. ruthenica* can reach 1.5 m, covering extensive grassland areas and effectively reducing soil water and heat loss (Wuriya et al., 1994). Its drought tolerance exceeds that of other leguminous forage species (Li et al., 2010). Therefore, *M. ruthenica* plays a vital role in grassland improvement, artificial grassland establishment, and soil-water conservation.

Owing to its rich morphological diversity, inferences regarding its phylogenetic position have remained controversial. When first recorded in 1753 (Linnaeus, 2015), *M. ruthenica* was classified as *Trigonella* because of its morphological characteristics, growth habit, and chromosome number similarities to *Trigonella* spp., and named *Trigonella ruthenica*. It was reclassified as *Medicago ruthenica* in 1841 (Small and Jomphe, 1989), changed to *Pocockia ruthenica* in 1872 as reported in the Wikimedia Species Database (<https://species.wikimedia.org/wiki/Trigonella>), renamed *Trigonella korshinskii* in 1945 and *Trigonella emodi* in 1955 (Latch, 1976), reclassified as *Melissitus ruthenicus* in 1976 and renamed *Meliotoides ruthenica* in 1982 (Soják, 1982). In the *Flora of China* (jointly published by China and the United States in 2010), *M. ruthenica* was reclassified as *Medicago* with the scientific name *Medicago ruthenica* (Wu et al., 2010). According to *Flora of China*, *M. ruthenica* includes one typical species and three varieties: *M. ruthenica* (*M. ruthenica* [L.] Soják), *Yinshan M. ruthenica* (*M. ruthenica* var. *inshanicus*), *Liaoxi M. ruthenica* (*M. ruthenica* var. *liaosiensis*), and *fine-leaved M. ruthenica* (*M. ruthenica* var. *oblongifolia*). Through integration of previous findings, the genetic separation between *M. sativa* and *M. ruthenica* may be more than that between *M. sativa* and *Medicago edgeworthii* (Campbell and Bauchan, 2002). However, precise classification of *M. ruthenica* requires additional molecular evidence.

Genome analysis has significant implications for crop phylogeny, morphology, and breeding. Chloroplasts and mitochondria are semiautonomous eukaryotic organelles that contain DNA. According to endosymbiotic theory, mitochondria and chloroplasts originated from free-living organisms, with an increasing number of genes being transferred to the nuclear genome over evolutionary time (Berry, 2003). These organelles contain essential genes for respiration and photosynthesis. In contrast to animal mitochondrial genomes, plant mitochondrial genomes exhibit remarkable variability and complexity in terms of genome size, gene content, gene arrangement, RNA editing, gene loss, and DNA sequence transfer, and are characterized by complex combinations, diverse structures, variable non-coding regions, and highly recombinogenic repetitive sequences (Kempken, 2011). Mutations in non-coding intergenic regions are repaired via non-homologous end-joining or break-

induced replication, resulting in genomic expansion, rearrangements, and chimerism. (Wang et al., 2024a). Repetitive sequences also influence the complexity of plant mitochondrial genomes, such as the presence or absence of inverted repeats were selected by species in changing environments (Christensen, 2013; Wu et al., 2015). *Medicago* species are usually considered as an inverted repeat (IR) lacking clade (IRLC) legume species that have been found newly regrown IR sequence during evolution (Choi et al., 2020). Plant mitochondrial genomes frequently undergo rearrangement and expansion: they can eliminate inaccurately repaired coding regions through natural selection, whereas inaccurately repaired non-coding regions tend to persist (Christensen, 2014). Back-mapping of long reads to these repeats verified the different conformations for each accession and that the varied conformations did indeed exist simultaneously. The existence of these various conformations and structural variations in *Morus* mitogenomes may influence their evolutionary dynamics, and their ability to adapt to environmental changes, and can also offer increased flexibility for interactions with the nuclear genome (Liu et al., 2024).

Besides, even among closely related species, mitochondrial genomes display substantial variation in size, structure, and genetic content (Kubo and Newton, 2008). In terms of synonymous substitution rate, compared to the rapidly evolving nuclear genome and the second most rapidly evolving chloroplast genome, mitogenomes evolve slowly in sequence (Wang et al., 2024b). Consequently, mitochondrial genome data have been widely applied in plant phylogenetic systematics and taxonomic studies (Ruprecht et al., 2017), including those on the gramineous forage *Elymus sibiricus*, legume forage *Medicago sativa* ssp. *sativa*, and medicinal plants such as *Carthamus tinctorius*, *Solanum tuberosum*, *Sorghum bicolor*, and *Glycine max* (Guisinger et al., 2010; Liu et al., 2021; Xiong et al., 2022; He et al., 2023; Wu et al., 2023; Zhang et al., 2023). Distinct types of mutation events in mitochondrial genomes enable species identification and evolutionary branch reconstruction across different timescales (Wang et al., 2021). Therefore, a comprehensive analysis of lentil (*Lens culinaris* Medik) mitochondrial genome can provide critical scientific evidence for refining its taxonomic placement.

The mitochondrial genomes of legume plants exhibit significant structural and functional diversity, providing abundant resources for studying the evolutionary relationships and genetic characteristics among these species. *M. ruthenica* is an important protein-rich forage crop belonging to the legume family. Studies of its nuclear and chloroplast genomes have revealed unique adaptive mechanisms and evolutionary traits under environmental pressure. The nuclear genome of *M. ruthenica* spans 903.56-904.13 Mb and contains approximately 50,162-50,268 protein-coding genes (PCGs). This gene number is larger than that in related species such as *Medicago truncatula* and *M. sativa* ssp. *sativa*, primarily due to repetitive element expansion (Yin et al., 2021). Phylogeographic analyses identified two distinct genetic lineages in *M. ruthenica* consistent with its east-west distribution. These lineages likely diverged during the Last Glacial Maximum and persisted in multiple refugia (Yan et al., 2021). Chloroplast genome phylogenies indicate a close relationship between *M. ruthenica*

and *M. truncatula* (Xie et al., 2021). However, cp genome annotations for *M. ruthenica* vary significantly, with reported gene counts of 107, 108, and 116 genes (Melton et al., 2015; Keller et al., 2017; Xu et al., 2021). Such discrepancies may arise from extensive structural rearrangements and gene losses in plant chloroplast genomes, potentially compromising the accuracy of phylogenetic inferences. The current taxonomic evidence for *M. ruthenica* remains inconclusive (Semerikov et al., 2022).

Mitochondrial genomes have been widely used to reconstruct phylogenetic relationships, shedding light on species' geographical distributions and migration histories (Arseneau et al., 2017). Mining functional sequence information is crucial for exploring plant phylogenies relationships. In the present study, we sequenced and annotated the mitochondrial genome of *M. ruthenica* and conducted comprehensive analyses focusing on genomic characteristics, repetitive sequences, and codon usage bias. Phylogenetic analysis was performed to clarify the genetic divergence in *M. ruthenica* in combination with existing reports on its breeding and plant research. Our findings provide a theoretical foundation for future research.

2 Materials and methods

2.1 DNA extraction, genome sequencing, and assembly

Fresh leaves of *M. ruthenica* were collected from Hohhot, Inner Mongolia, China (40.57°N, 111.93°E), and stored at the National Perennial Forage Germplasm Resource Nursery in Hohhot, China. Genomic DNA was isolated from fresh leaves using the Plant DNA Isolation Kit (Tiangen, Beijing, China). In this study, we first performed sequencing using the Novaseq6000 (Illumina, San Diego, CA, USA) and PromethION (Oxford Nanopore Technologies, Oxford, UK) platforms. For mitochondrial genome assembly, the Nanopore long-read data were initially aligned against reference gene sequences (https://github.com/xul962464/plant_mt_ref_gene) using minimap2 (v2.1) (Li et al., 2016). Subsequently, Canu (v2.0) was used for error correction and Bowtie2 (v2.3.5.1) was employed to align the second-generation data with the corrected sequence (Koren et al., 2017; Langmead and Salzberg, 2012). Finally, Unicycler (v0.4.8) with default parameters was employed to merge the matched second-generation data with the corrected third-generation data, aiming to obtain a complete and accurate mitochondrial genome sequence (Wick et al., 2015). Using Bandage (v0.8.1), we successfully assembled the mitochondrial genome of *M. ruthenica*. To ensure the reliability of the assembled mitochondrial genome, we conducted a comprehensive validation through coverage depth analysis using raw sequencing data from both NovaSeq 6000 and PromethION platforms (BioProject: PRJNA1291276; BioSample: SAMN49959260; SRA (Illumina): SRR34532195; SRA (ONT): SRR34532194). For the NovaSeq 6000 short-read data, we performed reference sequence alignment using Bowtie2 (v2.3.5.1)

(Langmead and Salzberg, 2012). In parallel, we processed Oxford Nanopore long-read data through reference-guided assembly with minimap2 (v2.1) (Li et al., 2010). Both alignment procedures were complemented by SAMtools (v1.9) for in-depth coverage analysis and quality assessment throughout the entire mitochondrial genome sequence (Li et al., 2009). Sequences of the *M. ruthenica* mitochondrial genome were deposited in the GenBank (accession number: PP723040).

2.2 Genome annotation

To annotate the PCGs in the *M. ruthenica* mitochondrial genome, we used the whole genomes of six previously published alfalfa species (*Medicago polymorpha* MW971561, *M. sativa* ssp. *sativa* NC_068105, *Medicago sativa* ssp. *falcata* OQ612687, *M. truncatula* NC_029641, *Medicago arabica* OX326972 and *Medicago lupulina* OY283147) as a reference. Genome annotation was performed using Geseq (v2.03) and PMGA (Li et al., 2025; Lowe and Eddy, 1997). tRNA and rRNA genes were verified with tRNAscan-SE (v2.0.12) using default parameters (Greiner et al., 2019), while open reading frames (ORFs) of ≥ 300 base pairs (bp) were identified via ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>). A circular mitochondrial genome map was generated using Organellar Genome DRAW (v1.3.1) (Beier et al., 2017).

2.3 Repeat sequence identification

The *M. ruthenica* mitogenome was analyzed for repeats-simple sequence repeats (SSRs, also called microsatellite) motifs using MISA (v2.1) (Benson, 1999), with detection thresholds set to mono-10, di-5, tri-4, tetra-3, penta-3, and hexa-3 repeats. Tandem Repeats Finder (v4.09) was applied with default parameters to identify tandem repeats (Kurtz, 2001), while dispersed repeats of ≥ 70 bp were characterized as forward, reverse, palindromic, and complementary repeats using REPuter (Iriarte et al., 2021). The analysis parameters included a Hamming distance of 3 and e-value cut-off of $1e-5$.

2.4 Codon usage bias analysis

Codon (v1.4.4) software was used to identify the codon usage of the mitochondrial genomes of *M. ruthenica*, *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina* (Peden, 2000). First, perl scripts were used to extract PCGs with a length greater than 300 bp, starting with ATG and terminating with TAA, TAG or TGA, while removing duplicate genes. Subsequently, the standard protein coding sequence (CDS) was input into Codon (v1.4.4) software in FASTA format, and the relative synonymous codon usage (RSCU) was calculated. The histogram was drawn using the R language ggplot2 package (Wickham, 2009).

2.5 Prediction of RNA editing sites

Based on six RNA-seq datasets of *M. ruthenica* deposited in the Sequence Read Archive (SRA, <https://www.ncbi.nlm.nih.gov/sra/>) under accession numbers SRR4140266, SRR4140268, SRR4140269, SRR4140270, SRR4140271, and SRR4140272, we established putative RNA editing sites in the mitochondrial PCGs. The alignment of RNA sequences for mitochondrial PCGs was performed using BWA (v0.7.15) (Li and Durbin, 2010). SNP identification was subsequently carried out with SAMtools (v1.17) and BCFtools (v1.17) (Li et al., 2009; Danecek et al., 2021). RNA editing sites were detected and annotated using REDO (v1.0) with its default parameters (Wu et al., 2018). To eliminate false editing sites, DNA-seq data were aligned to the *M. ruthenica* mitochondrial genome using BWA, BCF tools were used to identify genomic SNPs and exclude overlapping sites from RNA editing candidates.

2.6 Chloroplast-derived mitochondrial sequence identification

The chloroplast genome of *M. ruthenica* was *de novo* assembled using GetOrganelle (v1.7.5.3) software and performed annotation with PGA (Jin et al., 2020; Qu et al., 2019). Homologous fragments in the mitochondrial and chloroplast genomes were identified using BLASTn (v2.14.1+) (Chen et al., 2015), and gene transfer from the chloroplasts to mitochondria was visualized using TBtools-II (v2.136) (Chen et al., 2020).

2.7 Selective pressure calculation (Ka/Ks)

MAFFT (v7.526) was used for multiple sequence alignment of 26 common genes (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *cob*, *cox1*, *cox3*, *matR*, *matB*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*, *rpl16*, *rpl5*, *rps12*, *rps4*) between the mitogenomes of *M. ruthenica* and 33 other legume species (Katoh et al., 2019). Then import Ka/Ks Calculator (2.0) software (Wang et al., 2010), select the Yang-Nielsen (YN) model to calculate the non-synonymous substitution rate (Ka) and the synonymous substitution rate (Ks), and further calculate the Ka/Ks ratio. Finally, the Ka/Ks value box plot is drawn through the R language ggplot2 package.

2.8 Nucleotide diversity (Pi) analysis

Based on 30 common genes (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *ccmFn*, *cob*, *cox1*, *cox2*, *cox3*, *matR*, *matB*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*, *rpl16*, *rpl5*, *rps12*, *rps14*, *rps4* and *rps7*) in the mitochondrial genomes of *M. ruthenica*, *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina*. MAFFT (v7.526) was used for alignment, and then the alignment sequence was input into DnaSP (v6.12.03) for Pi calculation (Rozas et al., 2017).

2.9 Phylogenetic analysis

Phylogenetic analysis was performed using 18 shared PCGs (*atp4*, *atp8*, *ccmC*, *cob*, *cox1*, *cox3*, *matR*, *nad2*, *nad3*, *nad4*, *nad5*, *nad6*, *nad7*, *nad9*, *rpl16*, *rps12*, *rps3* and *rps4*) from 34 legume species mitogenomes. *A. speltoides* and *A. thaliana* were designated as the outgroup taxa. The common genes were extracted using the PhyloSuite (v1.2.3) software, and aligned with MAFFT (v7.526) (Zhang et al., 2020). Disordered positions and divergent regions were removed using Gblocks 0.91b (Talavera and Castresana, 2007). The Model Finder tool in PhyloSuite v1.2.3 (Kalyaanamoorthy et al., 2017; Zhang et al., 2020) was used to select optimal partitions for Bayesian and ML constructions. Subsequently, the MrBayes (v3.2.7) software was used to construct a Bayesian Inference (BI) tree (Ronquist et al., 2012), with the optimal model being GTR+F+I+G4. RAXML (v8.2.12) was employed to construct the maximum likelihood (ML) tree with 1000 bootstrap replicates (Stamatakis, 2006), and the optimal partition model was set as GTRGAMMAI.

2.10 Collinearity analysis

The mitochondrial genome sequences of seven species of *M. ruthenica* and *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina* were analyzed for covariance. The homologous sequences of these seven species were first aligned two-by-two based on BlastN (v2.13.0) to filter out the high-confidence homologous regions with a length of more than 300 bp (Chen et al., 2015). These segments were then analyzed for multiple covariance using MCScanX and plotted for multiple covariance mapping (Wang et al., 2012).

3 Results

3.1 Genomic characteristics of the mitochondrial genome in *M. ruthenica*

The *M. ruthenica* mitogenome assembly graph revealed differences in the distribution, connection, and coverage of repetitive sequences. Sites eight and nine harbored inversions, while site seven formed a macrocycle or two small subcycles (Figure 1). It was assembled into a single circular-mapping molecule of 354,988 bp (Figure 2). With a GC content of 45.13%, the mitogenome had a higher percentage than its AT content of 54.87% (Supplementary Table S1). Within the genome, the CDS spanned 29,844 bp, with 42.76% GC and 57.24% AT content (Supplementary Table S2). Transfer RNA (tRNA) genes constituted 1,309 bp with 51.57% GC and 48.43% AT contents, whereas ribosomal RNA (rRNA) genes spanned 5,239 bp with 51.67% GC and 48.33% AT contents. These results indicate that the GC content of protein-coding sequences are lower than that of tRNA genes.

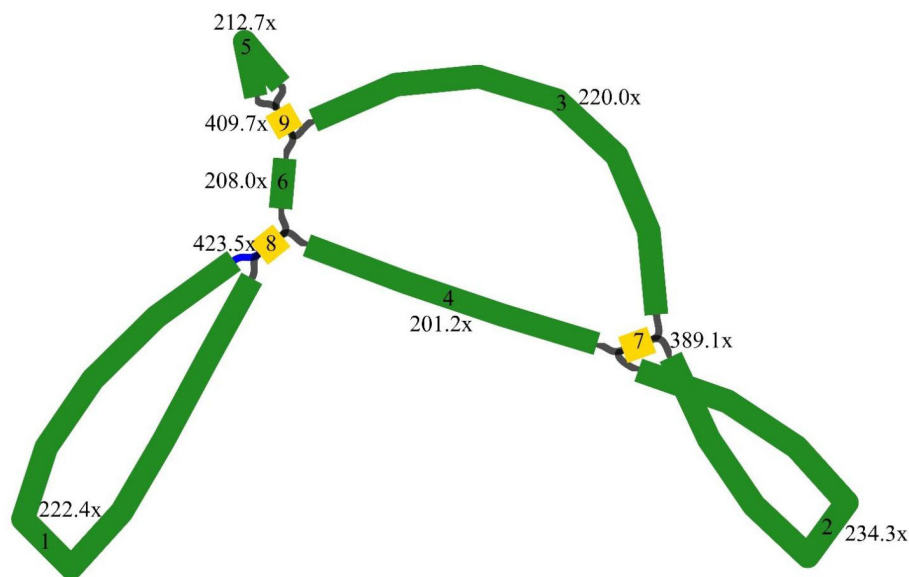


FIGURE 1

Genome assembly graph of the *Medicago* mitogenome. Nodes (yellow squares) represent genomic segments flanking repeats, edges (green segments) denote potential connections between these segments, the distinct coverage values of edges indicate differential connectivity patterns within repetitive regions.

Nad1, *nad2*, *nad5*, and *nad9* each contained four introns (Supplementary Table S3); *nad4* contained three; and the remaining genes lacked introns. Ribosomal protein genes spanned a total length of 5,994 bp; *rps1* was designated as a pseudogene, *rpl5* and *rps14* each had two copies, and only *rps10* and *rps3* contained one intron (Table 1). The three longest genes were *nad5* (2,010 bp), *matR* (1,986 bp), and *ccmFn* (1,737 bp), and the shortest was *atp9* (225 bp). A total of 18 tRNA and three rRNA genes were randomly interspersed among the PCGs. Not all genes were located on the positive strand; 12 tRNA genes and all rRNA genes were located on the negative strand. Only *trnA-TGC* and *trnT-TGT* contained one intron each, whereas *trnF-GAA* and *trnM-CAT* contained two or three copies.

3.2 Repetitive sequences in the *M. ruthenica* mitochondrial genome

Repeat elements such as SSRs often play crucial roles in genome variation, evolution, stability, and organization. In this study, repetitive sequence analysis of the *M. ruthenica* mitogenome was performed to explore its structural complexity. In total, 148 repeat sequences with lengths ranging from 29 to over 300 bp were identified in the mitochondrial genome, including 69 forward repeats (F) and 79 palindromic repeats (P) (Figure 3). Palindromic and forward repeats accounted for 53% and 47%, respectively. Specifically, 40 forward and 40 palindromic repeats were < 60 bp, 8 forward and 18 palindromic repeats measured 100–300 bp, and two forward and three palindromic repeats exceeded 300 bp. In contrast, the 90–99 bp range was the least frequent, containing only three palindromic repeats. This distribution

highlights preferential accumulation of small repeat elements in the *M. ruthenica* mitochondrial genome. In this study, 72 SSRs were identified in the mitochondrial genome of *M. ruthenica* through our analysis (Figure 4), revealing diverse compositions. These SSRs comprised nine mononucleotide repeats (12.50%), 24 dinucleotide repeats (33.33%), 15 trinucleotide repeats (20.83%), 21 tetranucleotide repeats (29.17%), two pentanucleotide repeats (2.78%), and one hexanucleotide repeat (1.39%).

Most small repeats were derived from nonfunctional regions of the *M. ruthenica* mitochondrial genome, such as the intergenic spacers (IGSs), totaling 742 bp in repeat length (Supplementary Table S4). Specifically, six SSRs were identified in the IGS between *atp1* and *atp6*, including dinucleotide (AG, GT, GA), trinucleotide (GGT), and tetranucleotide (TAAA, ATTA) repeats spanning 72 bp. Six SSRs were detected in the IGS between *trnE-UUC* and *trnH-GUG*, consisting of mononucleotide (T), dinucleotide (AG, TA), trinucleotide (AGT, CTT), and tetranucleotide (TAGA) repeats covering 72 bp. Two SSRs were found in the IGS between *trnT-UGU* and *nad1*, including mononucleotide (A), dinucleotide (TA, CT), and tetranucleotide (AAAT) repeats, totaling 42 bp. These three regions exhibited the highest densities of repeat types. In the intronic regions, five dinucleotide (TC) repeats were identified in *nad2* intron; *nad4* intron contained ten mononucleotide (C), five dinucleotide (AG), and four trinucleotide (AAG) repeats; and *nad5* intron harbored dinucleotide (AG) and trinucleotide (ATA) repeats spanning 64 bp. Only 2 SSRs were detected in the coding regions: three tetranucleotide (AATG) repeats in *rps3* and ten mononucleotide (A) repeats in *nad1*, totaling 22 bp.

A striking feature was the prevalence of A/T-rich SSRs (Supplementary Table S5), with 71 (98.61%) SSRs exhibiting A/T richness, including 24 SSRs composed exclusively of A/T bases.

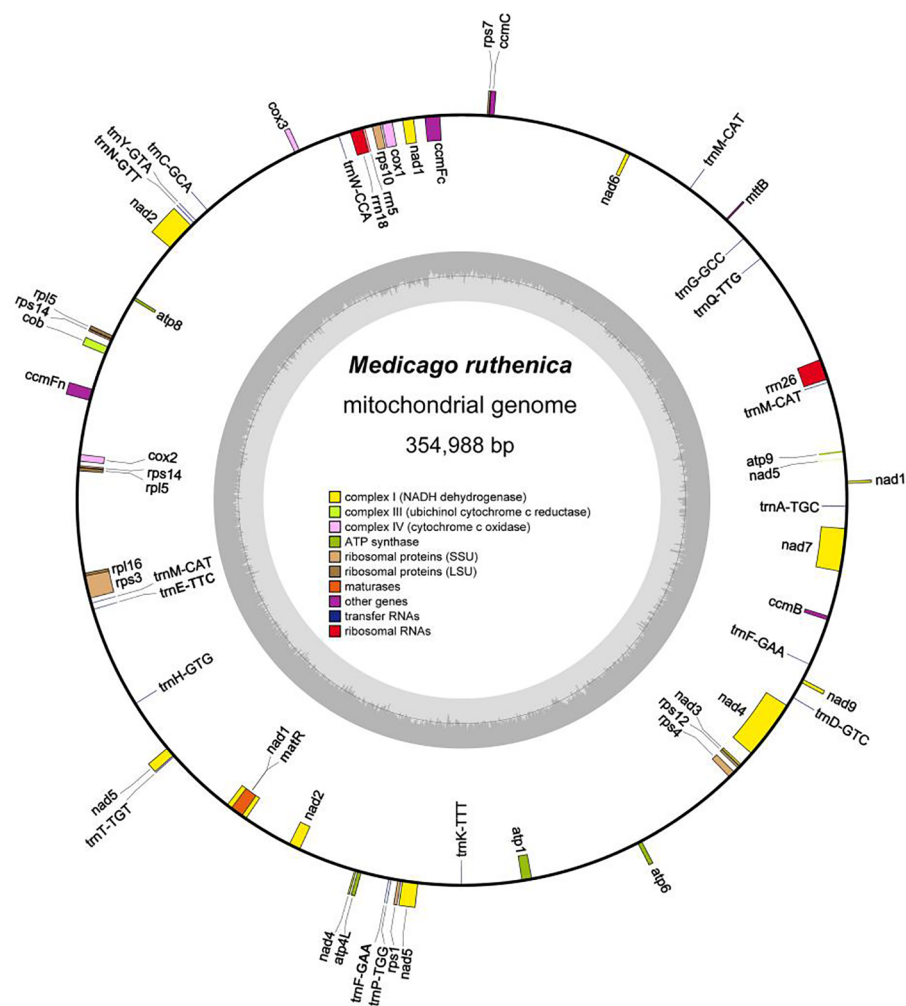


FIGURE 2

Circular diagram representing the complete mitochondrial genome of *M. ruthenica* and illustrating the gene organization. Genes positioned on the outer side of the circle are transcribed in the forward direction, while those on the inner side are transcribed in the reverse direction. The inner charcoal grey region indicates the GC content. Genes belonging to different functional categories are distinguished using various colors.

These included eight mononucleotide repeats (A/T), nine dinucleotide repeats (AT/TA), three trinucleotide repeats (AAT/TTA, TAT/ATA), and four tetranucleotide repeats (AAAT/ATTT, AATT/TTAA). The abundance of SSRs offers a rich resource for the development of molecular markers that are valuable for identifying and genetically characterizing *M. ruthenica*.

The *M. ruthenica* mitogenome contains 148 dispersed long repeat pairs (≥ 30 bp), including 69 forward and 79 palindromic repeats (Figure 5, Supplementary Table S6). Among these, 8 forward and 18 palindromic repeats measured 100–300 bp, one forward repeat spanned 400–500 bp, and three palindromic repeats and one forward repeat exceeded 3,000 bp. One pair of forward repeats and one pair of palindromic repeats are localized in the coding region (*atp1*, *cox2*; *matR*, *ccmFc*). One pair of palindromic repeats were localized to the tRNA genes, *trnF*-GAA and *trnF*-GAA. Three pairs of forward repeats and five pairs of palindromic repeats were localized to the introns. Four pairs of forward repeats and five pairs of palindromic repeats were localized in both the intron and

the IGS regions. We identified 18 pairs of forward repeats and 12 pairs of palindromic repeats localized in both the coding and the IGS regions, while the others were all found in the intergenic spacer regions. In addition, 11 pairs of forward repeats and 14 pairs of palindromic repeats were localized in the IGS regions between *atp1* and *atp6*. Moreover, 4 pairs of forward repeats and 12 pairs of palindromic repeats were localized in the IGS regions between *atp6* and *rps4*. Finally, three pairs of forward repeats and two pairs of palindromic repeats were localized in the IGS region between *nad2* and *atp8*.

Five tandem repeats (24–41 bp) were detected in the *M. ruthenica* mitogenome (Supplementary Table S7). One tandem repeat (CACCTGCAGTTTGGTGCAG) was located in the coding region of *ccmFn*. The remaining repeats resided in the IGS regions, including an 11 bp repeat (TTTCTTTTCT) between *nad6* and *ccmC*, 12 bp repeat (TACGTATATAAA) between *trnA*-TGC and *nad1*, 19 bp repeat (TTCCAGTTCTGTTATAATA) between *trnT*-UGU and *nad1*, and 21 bp repeat (ATAATTAGTGAACCTTATGAGA)

TABLE 1 Start and stop codon information of the *M. ruthenica* mitochondrial gene.

Gene Category	Group of genes	Gene name
Core genes	ATP synthase	<i>atp1, atp4, atp6, atp8, atp9</i>
	Cytochrome c biogenesis	<i>ccmB, ccmC, ccmFc*, ccmFn</i>
	Ubiquinol cytochrome c reductase	<i>cob</i>
	Cytochrome c oxidase	<i>cox1, cox2, cox3</i>
	Maturases	<i>matR</i>
Variable genes	Transport membrane protein	<i>mttB</i>
	NADH dehydrogenase	<i>nad1****, nad2****, nad3, nad4***, nad4L, nad5****, nad6, nad7****, nad9</i>
	Ribosomal proteins (LSU)	<i>rpl16, rpl5(2)</i>
	Ribosomal proteins (SSU)	<i>#rps1, rps10*, rps12, rps14(2), rps3*, rps4, rps7</i>
rRNA genes	Ribosomal RNAs	<i>rrn18, rrn26, rrn5</i>
tRNA genes	Transfer RNAs	<i>trnA-TGC*, trnC-GCA, trnD-GTC, trnE-TTC, trnF-GAA(2), trnG-GCC, trnH-GTG, trnK-TTT, trnM-CAT(3), trnN-GTT, trnP-TGG, trnQ-TTG, trnT-TGT*, trnW-CCA, trnY-GTA</i>

*: ntron number; #Gene: pseudo gene; Gene (2): number of copies of multi-copy genes.

between *rps7* and *ccmFc*. These insights into the mitogenome structural dynamics and repeats with potential functional implications enhance our understanding of *M. ruthenica* genome regulation and evolution.

3.3 Codon usage bias in PCGs

In this study, the codon usage characteristics of mitochondrial genomes of *M. ruthenica*, *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina* were analyzed (Figure 6, Supplementary Table S8). The results showed that all species followed the standard genetic framework of 21 amino acids corresponding to 64 codons. Methionine (Met) and Tryptophan (Trp) were only encoded by a single codon, while Arginine (Arg), Leucine (Leu) and Serine (Ser) had the most synonymous codons (six each). The RSCU analysis showed that 31 codons in *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. lupulina* and *M. sativa* ssp. *sativa* showed preference (RSCU > 1), 31 codons were non-preference (RSCU < 1), and codons AUG and UGG showed neutral selection characteristics (RSCU = 1). In *M. arabica*, 30 codons exhibited preference (RSCU > 1), 31 codons showed non-preference (RSCU < 1), while the codons UGG, UGA, and AUG displayed characteristics of neutral selection (RSCU = 1). It is worth noting that 30 codons of *M. ruthenica* showed preference (RSCU > 1), 32 codons were non-preference (RSCU < 1), and codon UGG and codon AUG showed neutral selection characteristics (RSCU = 1). The UGA terminator (Ter) of *M. ruthenica* showed preference selection characteristics (RSCU = 1.29), while the RSCU of UGA terminator (Ter) of other species showed non-preference (RSCU < 1), which was different from other species. In addition, both *M. ruthenica* and *M. polymorpha* have the largest RSCU value of 1.55, which is the codon UUA encoding Leu. The RSCU values of the codon UAA encoding Ter in *M. truncatula* and *M. sativa* ssp. *falcata* were the largest, which were 1.67 and 1.56, respectively. The codon GCU encoding Ala in *M. sativa* ssp. *sativa* and *M. arabica* had the largest RSCU value of 1.54. The codon GCU encoding Ala in *M. lupulina* had the largest RSCU value of 1.51. The number of codons in the genomes of *M. ruthenica*, *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina* were 9159, 9197, 9033, 8226, 9012, 8231 and 8732 respectively. Among the seven

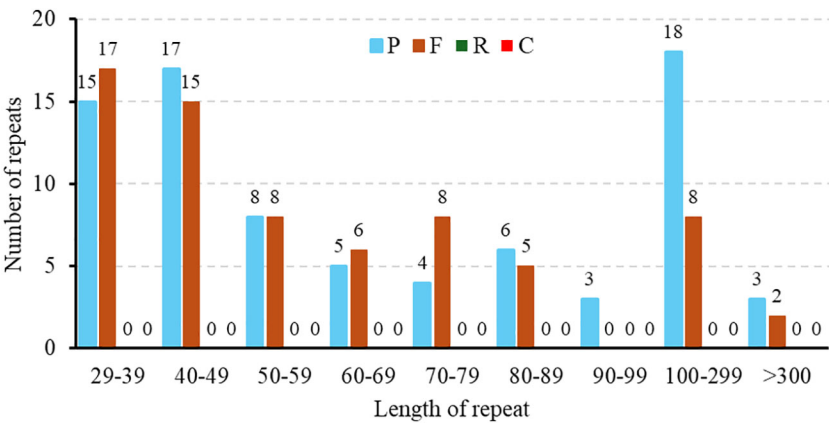


FIGURE 3 Distribution of dispersed repeat lengths in the *M. ruthenica* mitochondrial genome. The X-axis represents the categories of dispersed repeats, while the Y-axis shows the quantity of scattered repeats.



FIGURE 4

Patterns of SSR distribution in the *M. ruthenica* mitochondrial genome. The X-axis represents the different types of SSRs, while the Y-axis denotes the frequency of SSR repeats.

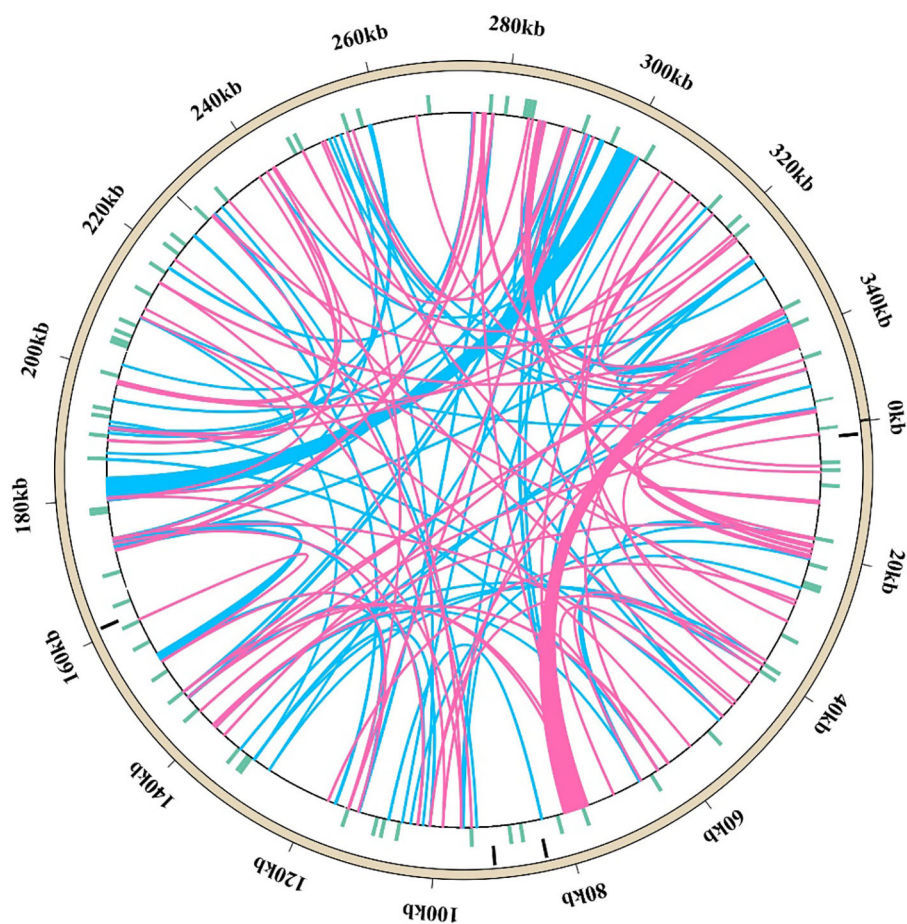


FIGURE 5

Arrangement of dispersed repeats within the mitochondrial genome of *M. ruthenica*. The outermost ring depicts the mitochondrial genome sequence, followed by layers representing SSR repeats, tandem repeats, and dispersed repeats (pink arcs indicate 69 forward repeats, while blue arcs represent 79 palindromic repeats).

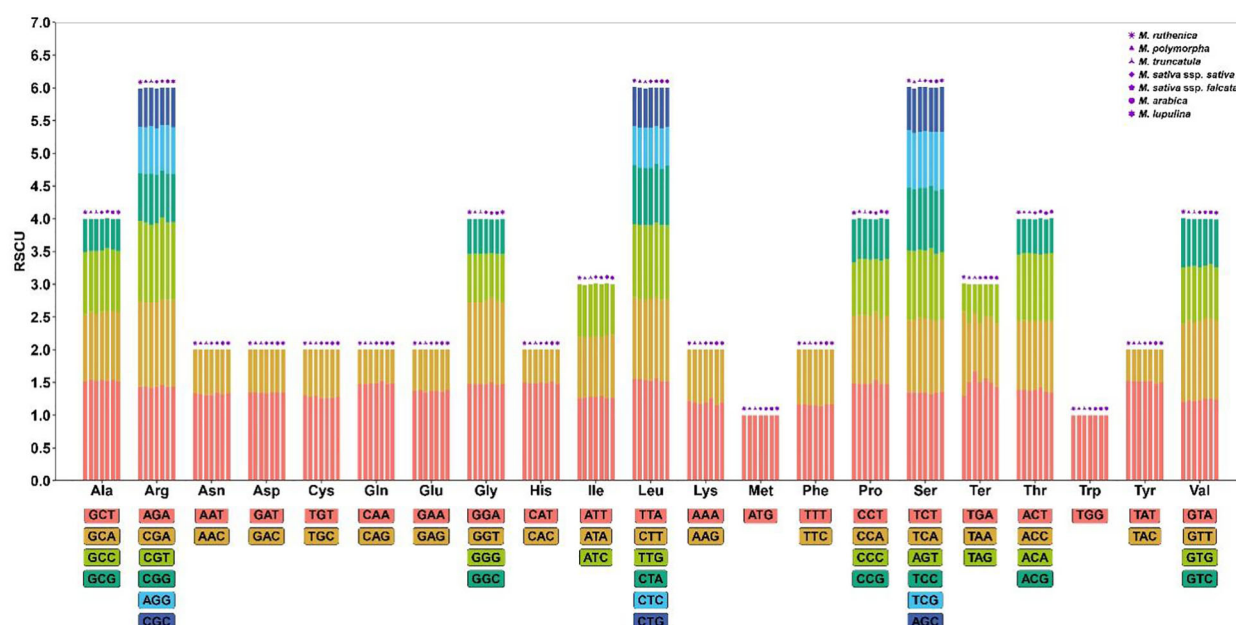


FIGURE 6
Relative synonymous codon usage (RSCU) of seven *Medicago* mitogenome. Codon families are on the X-axis. RSCU values are the number of times of a particular codon, relative to the number of times that the codon would be observed for a uniform synonymous codon usage.

species of *Medicago*, the number of codon UUU encoding Phe was the highest (257-341), and the number of codon UAG encoding Ter was the lowest (4-5).

3.4 Prediction of RNA editing sites

Mitochondrial RNA editing is an important post-transcriptional process that affects protein function and stability. A total of 207 RNA editing sites were identified in the 26 PCGs of the *M. ruthenica* mitochondrial genome (Figure 7). Among these PCGs, *ccmFc* exhibited the lowest editing frequency with only one editing site, whereas *nad2* was the most frequently edited gene with

24 sites. Both *nad7* and *cox2* displayed intermediate editing levels, with 19 sites each. Almost all RNA editing sites in the *M. ruthenica* mitogenome involved C to C-to-U conversion (Supplementary Table S9). Specifically, four C-to-A edits were detected in *cox2* and *nad6*, three G-to-U edits in *atp1* and *nad4L*, four A-to-G or G-to-A edits in *cox2* and *atp1*, one U-to-A edit in *atp1*, one C-to-G edit in *atp1*, and one G-to-C edit in *cox2*. Most RNA editing events occurred at the second nucleotide position (122 events), followed by the first position (71 events), with minimal editing at the third position (14 events). Among these, 100 converted hydrophilic amino acids to hydrophobic residues, 17 converted hydrophobic residues to hydrophilic residues, 18 maintained hydrophilicity, and 71 preserved hydrophobicity.

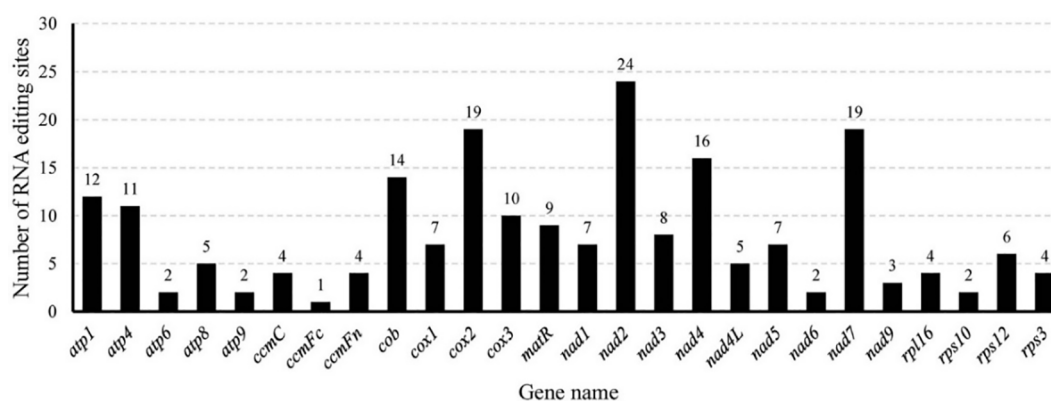


FIGURE 7
Localization of RNA editing sites in the mitochondrial PCGs of *M. ruthenica*.

In the *M. ruthenica* mitogenome, an editing event that converts CGA (Gln) to TAG (stop codon) was identified in *rpl16*. Three editing events converting ACG (Thr) to ATG (start codon) were detected in *cox2* and *nad7*, whereas only one event converting ATG (start codon) to ACG (Thr) was detected in *nad4L*. Among the amino acid substitutions, Pro to Leu was the most frequent (43 events), followed by Ser to Leu (39 events), and Ser to Phe (18 events), with other conversions being less common (Figure 8).

3.5 Chloroplast-originated sequences in the mitochondrial genome

The *M. ruthenica* mitochondrial genome spans 354,988 bp, approximately 2.80-fold larger than its chloroplast counterpart (126,939 bp). This size difference resulted in a more dispersed mitochondrial gene arrangement compared with that of the compact chloroplast genome (Figure 9). Analysis of chloroplast-specific (cp) sequences identified 42 fragments (46–3,546 bp) in the mitogenome, totaling 16,654 bp (4.69% of the mitochondrial genome) (Supplementary Table S10). These sequences migrate via diverse pathways: gene-to-gene, gene-to-IGS, IGS-to-IGS, and gene-to-intron/exon pathways. Among them, eight fragments derived from the chloroplast IGS regions were translocated to the mitochondrial IGS regions. Five sequences from chloroplast-encoded genes (*ndhF*, *petG*, *psaA*, *psaB*, *rbcL*) were transferred to the mitochondrial IGS regions, suggesting functional loss and integration into mitochondrial non-coding regions. Additionally, chloroplast *atpA* was converted to mitochondrial *atp1*. Nine migratory sequences from chloroplast tRNA genes (*trnA*-UGC, *trnD*-GUC, *trnE*-UUC, *trnF*-GAA, *trnH*-GUG, *trnN*-GUU, *trnP*-UGG, *trnM*-CAU, *trnQ*-UUG) were translocated to four mitochondrial IGS regions or five mitochondrial tRNA genes, accounting for 21.43% of all transferred sequences. We identified 12 migratory sequences from chloroplast ribosomal RNA genes (*rrn16* and *rrn23*) that were integrated into mitochondrial IGS regions.

3.6 Selective pressure analysis (Ka/Ks) in the mitochondrial genome

In this study, the Ka/Ks values of 26 common genes in the mitochondrial genomes of 34 legume species were analyzed to reveal the selection patterns of these genes during evolution (Figure 10, Supplementary Table S11). In general, the Ka/Ks values of *cox1*, *cob*, *nad7* and *nad4L* were lower than 0.3 in many species, showing a strong purifying selection signal, indicating that these genes were highly conserved in function. The Ka/Ks values of *atp8*, *atp4*, *rps4*, *nad4*, *ccmB*, *ccmFc* and *nad9* in multiple species were above one, showing obvious positive selection signals. The *ccmFc* showed significant positive selection pressure (Ka/Ks > 1) in *Arachis hypogaea*, *Lupinus albus*, *Apios americana*, *Glycyrrhiza glabra*, *Glycyrrhiza uralensis*, *Ghicine max*, *Gheine soja*, *Astragalus complanatus*, *Gleditsia sinensis*, *Oxytropis splendens*, *Oxytropis arctobia*, *Lotus japonicus* and *Lotus japonicus*. Among them, the Ka/Ks is the largest in *L. albus*, which is 1.53154. *A. hypogaea*, *L. albus*, *Leucaena trichanda*, *G. sinensis* and *Leucaena leucocephala* showed significant positive selection pressure on *atp8* (Ka/Ks > 1), and the Ka/Ks was the largest in *L. trichanda* and *L. leucocephala*, both 2.03829. In *L. japonicus*, *Senna tora*, *Senna occidentalis*, *L. trichanda* and *M. ruthenica*, *atp4* showed significant positive selection pressure (Ka/Ks > 1), and the Ka/Ks was the largest in *Senna occidentalis*, which was 1.35608. *Rps4* showed significant positive selection pressure (Ka/Ks > 1) in *L. albus*, *Vigna unguiculata*, *Vigna radiata*, *Milletia pinnata*, *G. max* and *G. soja*, and the Ka/Ks was the largest in *G. max* and *G. soja*, both of which were 1.63573. *Nad4* showed significant positive selection pressure (Ka/Ks > 1) in *G. glabra*, *G. uralensis* and *M. ruthenica*, and the Ka/Ks was the highest in *M. ruthenica*, which was 2.35843. *Nad9* showed significant positive selection pressure in *Ammopiptanthus mongolicus*, *Ammopiptanthus nanus*, *L. albus*, *A. americana*, *V. unguiculata*, *G. glabra*, *G. uralensis*, *V. radiata*, *G. max*, *G. soja*, *G. sinensis*, *A. complanatus*, *O. splendens*, *O. arctobia* (Ka/Ks > 1). Among them, the Ka/Ks was the highest in *O. splendens*, *A. complanatus* and *O. arctobia*, all of which were 2.04088. The

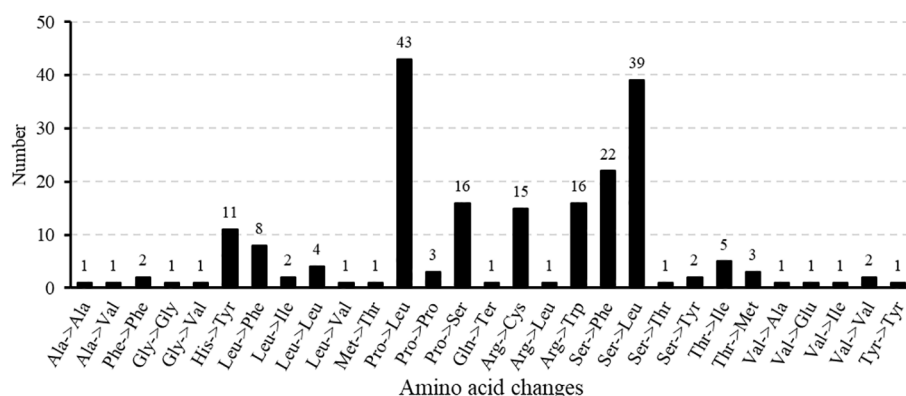


FIGURE 8
Occurrence of amino acid alterations induced by RNA editing in *M. ruthenica*.

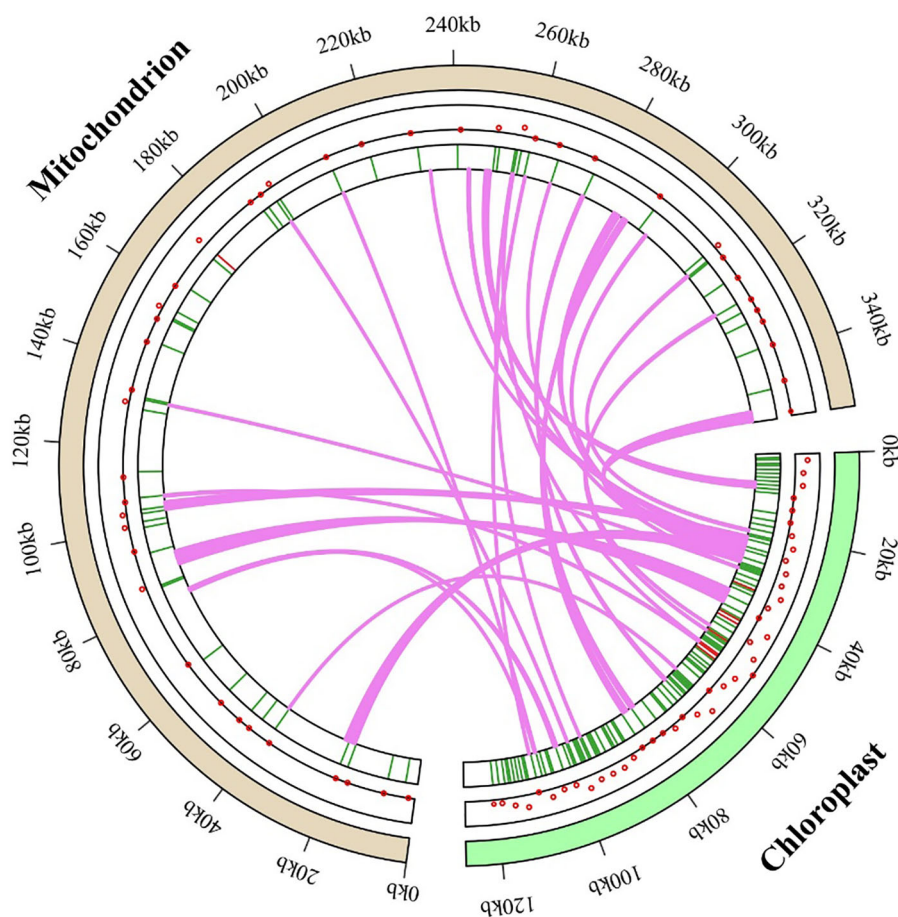


FIGURE 9
M. ruthenica genomic evolution, with gene transfer between chloroplast and mitochondrial genomes depicted using dots and heat maps within circular genome diagrams. Green segments represent the chloroplast genome, purple segments signify the mitochondrial genome, and lawn-green lines trace the pathways and integration sites of chloroplast-derived sequences into the mitochondrial genome. This visualization offers a comprehensive perspective on the dynamics of gene transfer between these organelles.

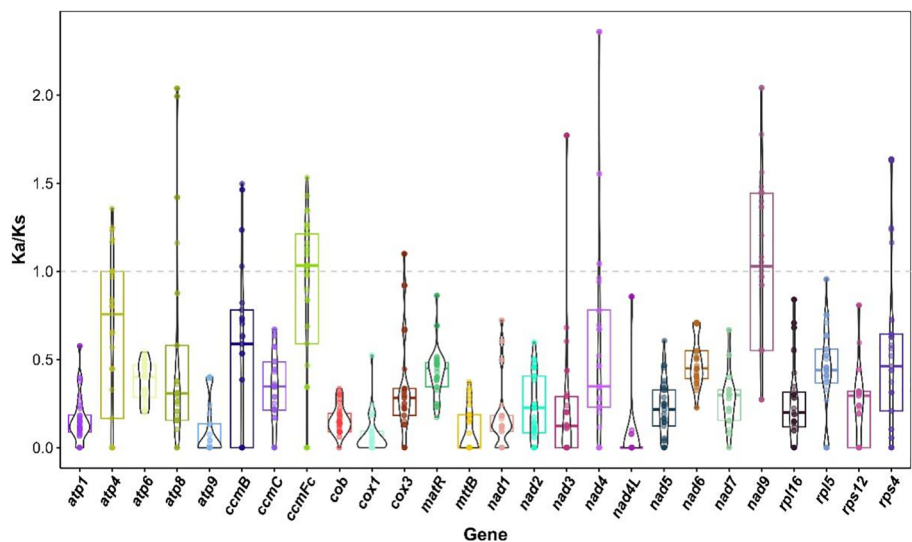


FIGURE 10
Ka/Ks values of 26 shared genes in mitogenomes of *M. ruthenica* versus 31 legumes species.

ccmB showed significant positive selection pressure ($Ka/Ks > 1$) in *A. mongolicus*, *A. nanus*, *L. albus*, *A. complanatus*, *O. splendens* and *O. arctobia*, and the Ka/Ks was the largest in *L. albus*, both of which were 1.49698.

In addition, the Ka/Ks values of *ccmFc*, *cox3*, *atp4*, *nad4*, *nad9* and *ccmB* in some species are close to the neutral evolution threshold ($Ka/Ks \approx 1$), suggesting that these genes may be in a state of transition from purification selection to positive selection. The Ka/Ks values of *ccmFc* in *V. unguiculata*, *G. uralensis*, *V. radiata*, *S. occidentalis*, *L. trichanda*, *Trifolium pratense*, *L. leucocephala* and *Pisum fulvum* ranged from 0.976155 to 1.09295. The Ka/Ks value of *cox3* in *P. fulvum* is 1.09941. The Ka/Ks values of *atp4* in *A. hypogaea*, *G. glabra*, *G. uralensis*, *A. complanatus*, *L. leucocephala*, *O. splendens*, *O. arctobia* ranged from 0.967073 to 1.00096. The Ka/Ks values of *nad4* in *G. max*, *L. trichanda* and *G. soja* ranged from 0.96150 to 1.04318. The Ka/Ks values of *nad9* in *A. hypogaea*, *L. japonicus*, *S. tora* and *S. occidentalis* ranged from 0.968876 to 1.08186. The Ka/Ks value of *ccmB* in *G. sinensis* was 1.02828. It is worth noting that the Ka/Ks values of multiple genes in seven *Medicago* species were zero or not applicable. Through the above analysis, this study not only comprehensively covers the changes of Ka/Ks different genes in various species, but also highlights the importance of key genes in evolution, which brings new clues to explore the evolutionary mechanism of the mitochondrial genome of legume plants.

3.7 Nucleotide diversity analysis (Pi) in the mitochondrial genome

In this study, the nucleotide diversity index (Pi) of 30 common genes in seven *Medicago* species was analyzed to reveal their genetic variation characteristics (Figure 11, Supplementary Table S12). Some genes showed high nucleotide diversity, for example, *atp6* (0.00632), *atp8* (0.00434), *cox2* (0.00613), *matR* (0.00235), *rpl5* (0.00205), *rps14*(0.00377), their Pi were greater than 0.002, these genes may have experienced relatively loose selection pressure or adaptive evolution. On the contrary, the Pi values of *atp4*, *atp9*, *mttB*, *nad3*, *nad4L* and *rps12* were zero, indicating that these genes were highly conserved in legumes. These results not only show the

functional differentiation pattern of evolutionary conserved regions and variable regions in the mitochondrial genome of legumes, but also provide important clues for understanding the adaptive evolution mechanism of the genus. By comparing the nucleotide diversity of different genes, we can further explore the role of mitochondrial genome in species evolution and its potential biological significance.

3.8 Phylogenetic analysis

In this study, based on 18 mitochondrial common genes (*atp4*, *atp8*, *ccmC*, *cob*, *cox1*, *cox3*, *matR*, *nad2*, *nad3*, *nad4*, *nad5*, *nad6*, *nad7*, *nad9*, *rpl16*, *rps12*, *rps3* and *rps4*), the phylogenetic trees of 34 legume species were constructed to determine the evolutionary status of *M. truncatula*. *A. speltoides* and *A. thaliana* served as outgroups. Our results reveal identical topology in the phylogenetic trees constructed by RAxML (ML) and MrBayes (BI) analyses (Figure 12). Maximum Likelihood (ML) analysis showed that about 88.24% (30/34) of the nodes showed strong support (bootstrap value $\geq 70\%$), of which 27 nodes (79.41%) reached full support (bootstrap value = 100%). Bayesian inference (BI) results further verify the topological reliability, and 91.18% of the nodes (31/34) have a posterior probability ≥ 0.95 . In addition, the phylogenetic tree topology supports the differentiation of the 36 species in this study into ten branches.

The first clade is composed of *Medicago* and *Trifolium*. Seven species of *Medicago* (*M. ruthenica*, *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina*) are clustered into a small clade, and three species of *Trifolium* (*T. aureum*, *T. meduseum* and *T. pratense*) are clustered into another small clade. The second largest clade consists of two species of *Pisum* (*P. fulvum* and *P. sativum*). The third branch is composed of *Astragalus* and *Oxytropis*. Among them, *A. complanatus* is a small branch, *Oxytropis splendens* and *O. arctobi* are clustered into another small branch. The fourth clade consists of two species of *Glycyrrhiza* (*G. uralensi* and *G. glabra*). The fifth branch consists of two species of *Lotus* (*L. corniculatu* and *L. japonicus*). The sixth branch was composed of *Milletti*, *Apios*, *Vigna* and *Glycine*, in which *M. pinnata* and *A. americana* were

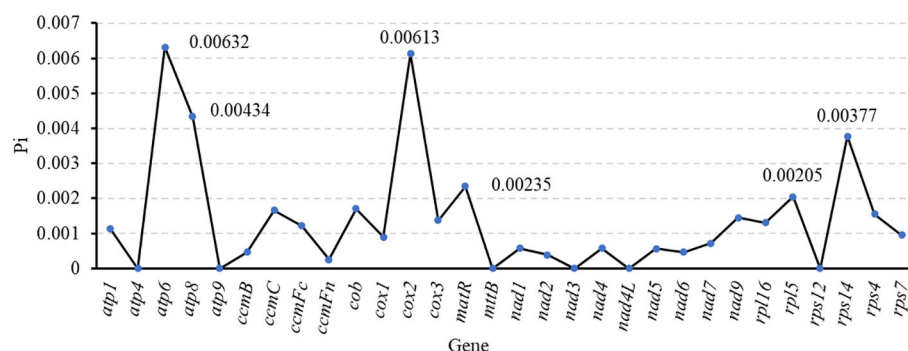
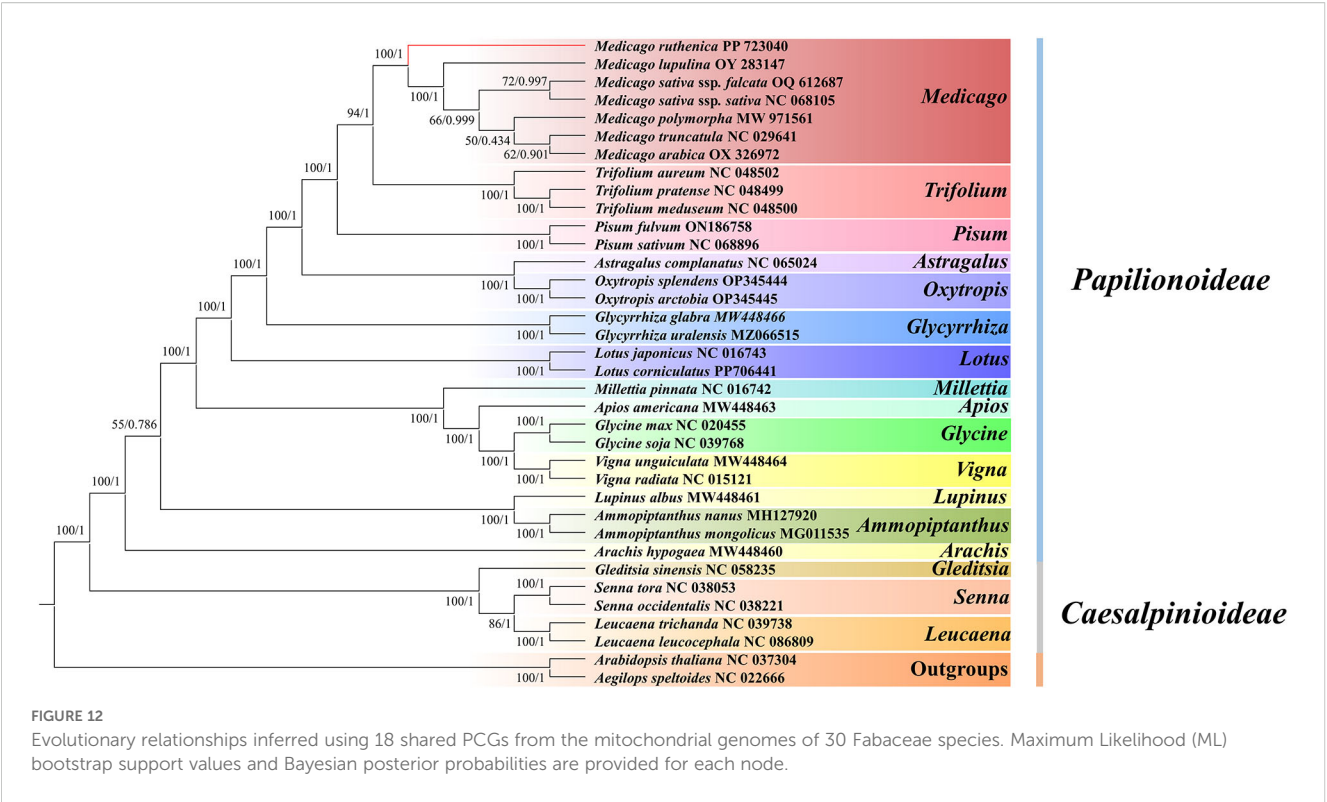


FIGURE 11
Nucleotide diversity of seven *Medicago* species (30 shared genes).

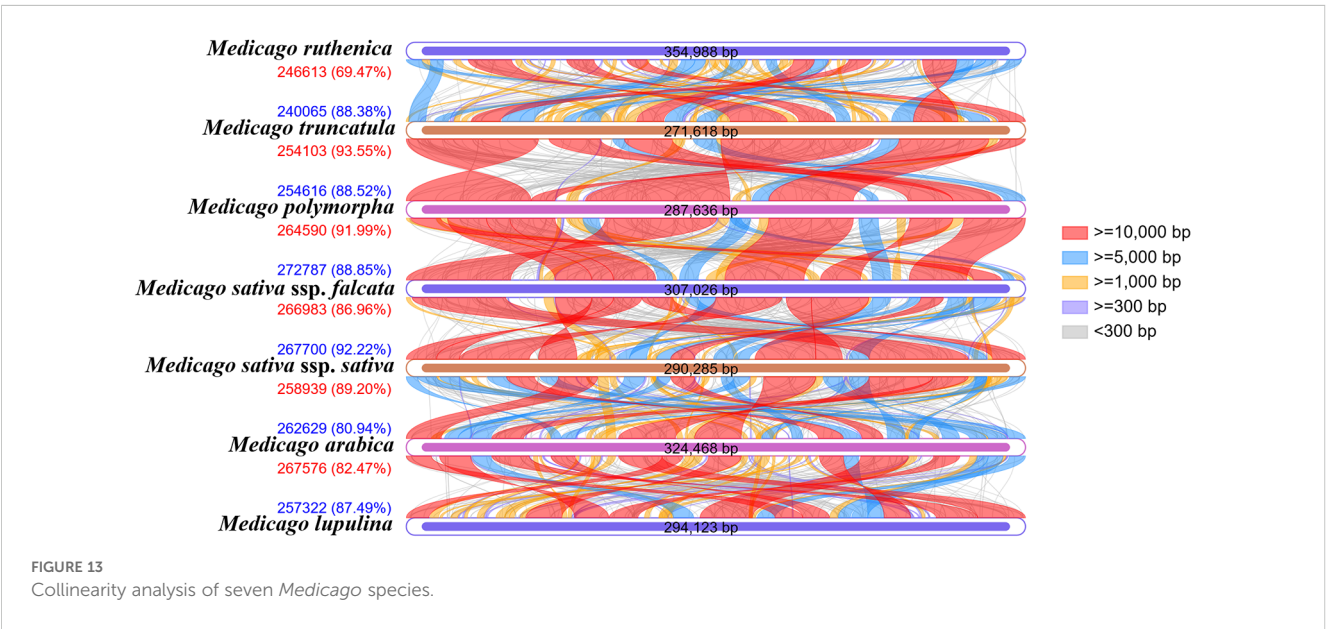


separated into a small branch, *V. radiata* and *V. unguiculata* were clustered into a small branch, and *G. uralensis* and *G. glabra* were clustered into another small branch. The seventh branch is composed of *Lupinus* and *Ammopiptanthus*, in which *L. albus* is a small branch alone, *A. mongolicus* and *A. nanus* are clustered into another small branch. The eighth branch is composed of *A. hypogaea* of *Arachis*. The ninth branch was composed of *Gleditsia*, *senna* and *Leucaena*, in which *G. sinensis* was separated into a small branch, *S. tora* and *S. occidentalis* were clustered into a small branch, and *L. trichanda* and *L. leucocephala* were clustered

into another small branch. The tenth clade is the outgroup, which is composed of *A. thaliana* of *Brassicaceae* and *A. speltoides* of *Brassicaceae*.

3.9 Collinearity analysis

Collinearity analysis reveals the evolutionary associations between species by exploring the conservation of homologous gene sequences (Figure 13, Supplementary Table S13). In this



study, the mitochondrial genomes of the seven *Medicago* species were compared pairwise. The length of homologous regions between *M. truncatula* and *M. polymorpha* was the largest, at 254,103 bp (93.55% of its mitochondrial genome) and 254,616 bp (88.52%), respectively, indicating that the two species had high similarity and conservation at the genomic level. Further analysis showed that the homologous regions of *M. polymorpha* and *M. sativa* ssp. *falcata* also accounted for 91.99% and 88.85%, respectively, also indicating high similarity. The homologous regions of *M. sativa* ssp. *falcata* and *M. sativa* ssp. *sativa* accounted for 86.96% and 92.22%, respectively. The homologous regions of *M. sativa* ssp. *sativa* and *M. arabica* accounted for 89.20% and 80.94%, respectively. The homologous regions of *M. arabica* and *M. lupulina* accounted for 82.47% and 87.49%, respectively. The coverage of homologous regions between *M. ruthenica* and *M. truncatula* was the lowest, 69.47% and 88.38%, respectively, suggesting that there were differences in genomic structure between them. In summary, *M. truncatula* is most closely related to *M. polymorpha*, whereas *M. ruthenica* is relatively distantly related to *M. arabica*. These results are consistent with the phylogenetic relationship of the seven *Medicago* species as inferred from the phylogenetic tree.

4 Discussion

Mitochondria, as essential semiautonomous organelles harboring their own genetic material in most eukaryotes, encode certain RNAs and polypeptides, and typically function in energy metabolism, development, and reproduction (Dyall et al., 2004; Kuzmin et al., 2005). Despite continuous advancements in sequencing technologies, plant mitochondrial genomes are increasingly variable and diverse. The size, structure, and gene content of mitochondrial genomes vary dramatically across angiosperms. The complexity of assembling plant mitochondrial genomes has led to a lag in the study of plant mitochondrial genomes (Knoop, 2013). In this study, we report the basic characteristics of the *M. ruthenica* mitochondrial genome for the first time. Our findings provide a valuable foundation for understanding the genomic function, inheritance, and evolutionary trajectory of *M. ruthenica*.

4.1 Mitochondrial genome analysis of *M. ruthenica*

We obtained a circular mitochondrial genome of 354,988 bp in length, which is larger than those of *M. truncatula* (271,618 bp), *M. polymorpha* (287,636 bp), *M. sativa* ssp. *falcata* (307,026 bp), *M. sativa* ssp. *sativa* (290,285 bp), *M. arabica* (324,468 bp) and *M. lupulina* (294,123 bp), but similar to those of *T. meduseum* (348,724 bp) and *O. splendens* (361,579 bp). The expansion of non-coding regions may have contributed to the evolution of larger and more complex plant mitochondrial genomes (Goremykin et al., 2012). A high guanine-cytosine (GC) content is common in plants (Smith

and Keeling, 2015), reported GC contents of legume mitochondrial genomes range from 44.67% to 45.45%. Notably, the *M. ruthenica* mitogenome exhibits a GC content of 45.13%, which falls within the GC content range (44–46%) of the related species of *Medicago*, including *M. truncatula* (45.39%), *M. polymorpha* (45.26%), *M. sativa* ssp. *falcata* (45.11%), *M. sativa* ssp. *sativa* (45.34%), *M. arabica* (44.92%) and *M. lupulina* (45.36%). Generally, plants species with a GC-biased mitochondrial genome also GC-rich in plastid genome (Smith and Keeling, 2015). GC base pairs with three hydrogen bonds were stronger than AT base pairs (two hydrogen bonds), resulting in greater stability of the *M. ruthenica* mitochondrial genome sequence. Higher GC content generally correlates with increased DNA stability (Chen and Skylaris, 2021). High GC content is often associated with functional advantages, such as stable gene expression and regulatory roles, which may be favored by natural selection (Muyle et al., 2011). Conversely, elevated GC content can lead to increased mutation and recombination rates, potentially enhancing genetic diversity and adaptability to changing or extreme environments (Kiktev et al., 2018).

Codon usage bias is of great significance in the mitochondrial genome. It not only affects the efficiency and accuracy of protein synthesis, but also has a close relationship with the adaptability and evolution of organisms (Jia and Higgs, 2008). This study systematically revealed the evolutionary characteristics of codon bias in the mitochondrial genomes of *M. ruthenica* and *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina*. The results showed that the seven *Medicago* species exhibited significant codon usage conservation while maintaining the standard genetic coding framework. Among them, the types and quantities of preferred codons (RSCU > 1) of *M. truncatula*, *M. polymorpha*, *Medicago sativa* ssp. *falcata*, *M. sativa* ssp. *sativa* and *M. lupulina* were completely consistent. Notably the UGA terminator codon (Ter) of *M. ruthenica* indicated a preferential usage (RSCU = 1.29), while the UGA terminators of other species showed a non-preferential usage (RSCU ≤ 1). This difference may indicate the fine-tuning of the mitochondrial genome of *M. ruthenica* in the termination codon recognition mechanism, such as the regulation of tRNA abundance or the adaptive adjustment of translation termination efficiency. This species-specific feature may be related to its unique ecological adaptability, but the underlying molecular mechanism requires verification by tRNA expression profile and translation elongation kinetics experiments (Wu C. et al., 2019). We found that most PCGs in *M. ruthenica* exhibited a codon usage bias toward A/U-ending codons, which is consistent with the A/U preference observed in the chloroplast genome (Unsel'd et al., 1997). Monocotyledonous plants such as *Zea mays* typically show an enrichment of C/G-ending codons, whereas dicotyledonous plants such as *Taraxacum* species, have higher proportions of A/U-ending codons (Sharma et al., 2019; Tian et al., 2021; Fennoy and Bailey-serres, 1993). This A/U preference may reflect evolutionary adaptation, as A/U-rich codons dominate early genetic code evolution and could be shaped by amino acid functional spectra or gene expression optimization. However, natural selection and

mutational pressures have also been recognized as significant drivers of codon usage bias (Parvathy et al., 2022; Yang et al., 2024).

This finding provides a novel entry point for the subsequent study of the association between mitochondrial genome evolution and species differentiation. In addition, only seven closely related species of *Medicago* are included as research samples, and the sample range is narrow. The limited number of samples may not fully and accurately reflect the overall pattern of mitochondrial genome codons in the entire *Medicago* species. Further studies are urgently needed to further expand the sample range, covering more *Medicago* species with different ecotypes and geographical distribution, as well as other plant groups with different genetic relationships with *Medicago*, and to carry out more extensive and in-depth comparative analysis, so as to fully verify the generalizability of the patterns identified herein and provide a more robust foundation for further understanding the theory of plant mitochondrial genome evolution.

4.2 Mitogenome repeats of *M. ruthenica*

Repeat sequences are critical factors in mitogenomic structural variation. The variations of mitogenome size and structure are thought to be caused by repeat-mediated duplication and recombination (Liu et al., 2024). We identified 148 dispersed repeats in the mitochondrial genome, with no inverted or complementary repeats, a pattern observed in the *M. ruthenica* chloroplast genome. Mitogenome copy numbers may predict plant evolutionary rates, the lower copy numbers may provide fewer templates, homologous recombination may act efficiently with a higher copy number (Zwonitzer et al., 2024). These 148 repeats belong to the imperfect repeats, are prevalent in mitogenomes, and are characterized by short, evolutionarily dynamic sequences (Fontecilla-Camps, 2023). Short insertional repeats are abundant in mitogenomes owing to their flexible nature and susceptibility to breaks, gene recombination, and conversion during plant evolution (Liu et al., 2024). Whereas the higher abundance of long repeats indicates that the mitochondrial genome may undergo more recombination events (Wang et al., 2025). In contrast, a lower abundance of long repeats or a higher abundance of short repeats in *M. ruthenica* suggests that its mitochondrial genome may undergo fewer recombination events during evolution. These short repeats are particularly frequent and large in vascular plants, potentially signaling adaptive strategies for their survival and evolution (Shakyawar et al., 2009). The presence or absence of inverted repeats significantly affects mitochondrial genome stability, replication, and evolution. Angiosperms such as *Silene noctiflora* exhibit enhanced genomic stability owing to the lack of inverted repeats (Kanazawa et al., 1998). Conversely, inverted repeats in other species facilitate complex rearrangements and influence replication mechanisms (Wynn and Christensen, 2019). Recent reports have indicated that the yellow-flowered *M. ruthenica* harbors large inverted repeats in its chloroplast genome, suggesting evolutionary novelty (Wu et al., 2015). Similarly, *M. minima*, despite being an ancestral member of the IRLC, lost a 9 kb

inverted repeat, which may explain differences in reproductive modes and environmental adaptation during evolution (Choi et al., 2020). Understanding these structural elements is crucial for unraveling the evolutionary complexity of mitogenomes across plant species.

The SSRs contribute to genomic evolution and diversity through rapid evolutionary changes associated with high mutation rates (Čechová et al., 2018). SSRs are classified as mono-, di-, tri-, tetra-, penta-, or hexanucleotide repeats. In this study, 72 SSRs were identified in the *M. ruthenica* mitogenome, accounting for 1.10% of its mitochondrial genome, a proportion relatively close to that of *M. truncatula* (2.41%) (Wen et al., 2024). Dinucleotide (33.33%) and tetranucleotide (29.17%) repeats predominated, with A/T bases constituting 98.61% of all SSRs. Mononucleotide and dinucleotide repeats are common in legume species, including *Ormosia boluoensis*, *A. mongolicus*, and *G. sinensis* (Li et al., 2018; Kumar et al., 2020; Liu et al., 2021; Xie et al., 2021), whereas tetranucleotide and hexanucleotide repeats are rare in plant mitogenomes, with the exception of *Trigonella foenum-graecum* (Yu et al., 2018; Yang et al., 2021). A/T- or TA/AT-rich repeats are also prevalent in legumes, such as *M. truncatula*, *T. foenum-graecum*, and *G. sinensis* (Li et al., 2018; Kumar et al., 2020; Liu et al., 2021). Most SSRs were distributed in the IGS regions, a pattern observed in other plants, such as *E. sibiricus* and *M. polymorpha* (Choi et al., 2020; Zhang et al., 2023). SSRs have potential as molecular phylogenetic markers because of their variability and distribution; however, their application in phylogenetic studies is challenged by the nonlinear dynamics in mitochondrial genomes (Choi et al., 2020).

4.3 RNA editing events and number of *M. ruthenica*

Mitochondrial rRNAs and tRNAs play crucial roles in cellular stress responses and resistance mechanisms as mitochondrial DNA acts as a stress sentinel that can activate nuclear DNA damage and repair pathways, while rRNAs and tRNAs are involved in post-transcriptional processing (Zhao et al., 2016). The *M. ruthenica* mitogenome contains three rRNAs genes (*rrn18*, *rrn26*, and *rrn5*), a feature shared with most *Medicago* species; however, its 18 tRNA genes exceed the numbers reported in Zhongmu No.1, Zhongmu No.4, *M. minima*, and *M. truncatula* (Nishikawa et al., 2005; Choi et al., 2020; Kumar et al., 2020; Liu et al., 2021). The relationship between tRNA gene numbers in mitochondrial genomes and resistance mechanisms is complex, as approximately 25% of tRNAs exist in organellar genomes (Wu Z. et al., 2019). These tRNAs may function as regulators of stress responses, by generating stress-induced tRNA-derived fragments (Matsui et al., 2013), which could enhance *M. ruthenica* tolerance to abiotic stress.

The post-transcriptional modification of RNA molecules, known as mitochondrial RNA editing, alters their nucleotide sequences and plays a vital role in regulating plant stress resistance. Our results indicate that C-to-U editing occurs in the *M. ruthenica* mitochondrial genome, a phenomenon conserved

across plant species such as *A. thaliana*, *Triticum aestivum* and *Poa pratensis* (Sablok et al., 2017; Hu et al., 2021; Ai et al., 2024). Editing is essential for accurate mitochondrial protein coding, and is influenced by selective pressures. C-to-U editing modifies codons to increase protein hydrophobicity; these hydrophobic proteins are typically involved in stress responses and signaling pathways (Gualberto et al., 1991; Giegé and Brennicke, 1999). RNA editing rates can be modulated by environmental stresses; for example, C-to-U editing decreases under heat or cold stress in *Arabidopsis* but increases under alkaline stress in *Oryza sativa* (Gosslau et al., 2001). Understanding these RNA editing mechanisms could facilitate the development of crops with enhanced stress tolerance.

4.4 Ka/Ks and Pi analysis of *M. ruthenica*

The Ka/Ks value, the ratio of non-synonymous to synonymous substitutions, is often used to evaluate the evolutionary rate of genes under selection pressure (Park et al., 2018). When Ka/Ks > 1, it indicates that the gene is under purified selection, suggesting that the gene has undergone adaptive changes in evolution, which may be related to its functional adjustment under specific environmental pressures (Yang and Bielawski, 2000). In this study, the dynamic role of natural selection pressure in the evolution of extranuclear genomes was revealed by analyzing the Ka/Ks values of 26 common genes in the mitochondrial genomes of 34 legume species. The Ka/Ks values of most genes (such as *cox1*, *cob*, *nad7* and *nad4L*) were significantly lower than 0.3, indicating that these genes were subjected to strong purification selection pressure. These genes participate in the electron transport, proton pump and ATP synthesis of the mitochondrial electron transport chain by encoding the key subunits of the respiratory chain complex. Their highly conserved sequence features are likely essential for maintaining the core function of mitochondrial energy metabolism (Sloan et al., 2012). In this study, some genes showed significant positive selection signals (Ka/Ks > 1) in specific lineages, and the functional optimization of these genes may directly enhance the adaptability of plants to adversity (Begum and Mondal, 2021). The Ka/Ks values of some genes were close to the neutral threshold (0.976155–1.09295), suggesting that they may be in the transition stage from purified selection to positive selection (Yang et al., 2023). For example, the *ccmFc* and *cox3* genes of *P. fulvum*, the *nad4* gene of *G. max*, the *nad9* gene of *A. hypogaea* and the *ccmB* gene of *G. sinensis* all showed such characteristics. This dynamic change provides a key clue for the study of selection pressure conversion, which is helpful to understand the evolution of gene function under different environmental pressures. Plants species with high copy numbers result in relatively slow “normal” substitution rates of mitochondrial genomes in most plant species. Lower copy numbers lead to various evolutionary rates, with some species experiencing rate acceleration having relatively lower copy numbers, while others with low copy numbers maintain low substitution rates without experiencing rate acceleration (Zwonitzer et al., 2024; Wang et al., 2024b).

Nucleotide diversity index (Pi) is one of the important means to measure the genetic diversity of species. It reflects the degree of genetic variation within the population by calculating the average nucleotide difference between each pair of sequences in a specific gene region (Konopiński, 2022). The high Pi usually means that the species has rich genetic diversity, which is of great significance for their adaptation to environmental changes, resistance to diseases and long-term evolution (Massa et al., 2024). In botanical studies, the PCGs in the mitochondrial genome are considered to be ideal for assessing genetic diversity due to their conserved and easy analyze characteristics (Wang et al., 2023). Among the seven *Medicago* species studied, some genes showed relatively high Pi values (ranging from 0.00205 to 0.00613), which may reveal the adaptive fine-tuning or relaxation of environmental selective pressure experienced by these genes during evolution (Mousavi-Derazmahalleh et al., 2018). Specifically, oxidoreductase-related genes (*cox2*) and ATP synthase-related genes (*atp6*) showed significant genetic diversity. Among them, the Pi value of *cox2* gene is as high as 0.00632, and the Pi value of *atp6* gene is 0.008724. This genetic variation may enable *M. ruthenica* to more effectively regulate its energy metabolism processes in different environments to adapt to specific ecological conditions. In addition, the ribosomal protein gene *rps14* also showed high genetic diversity (Pi = 0.00377). As the key structure of intracellular protein synthesis, ribosome function depends on the involvement of various ribosomal proteins. The ribosomal protein encoded by *rps14* gene plays an important role in the ribosome assembly. The variation of *rps14* gene may optimize the rate of ribosome assembly, so that *M. ruthenica* can respond more flexibly to changes in translation needs (Figueroa et al., 1999). This mechanism may help plants quickly adjust the rate of protein synthesis to maintain normal physiological functions in the face of environmental stress. Taken together, the high genetic diversity of these genes may reflect the adaptive adjustment strategies of *Medicago* plants in specific habitats. Therefore, in-depth exploration of the functions of these genes and their interactions with the environment can not only help to reveal the evolutionary history of *Medicago* plants, but also provide key scientific support for future conservation strategies and breeding practices.

4.5 Phylogenetic relationship of *M. ruthenica*

Since the establishment of *Medicago*, this genus has become an important research direction in botany. To date, the reported taxonomic studies for *M. ruthenica* have employed methods such as SSR/ISSR analysis, ITS2 sequence analysis at the molecular level (Takezawa and Minami, 2004), flavonoid compound type analysis at the chemical level, pollen morphology or karyotype analysis at the cellular level (Rehman et al., 2022), and chloroplast genome analysis at the organelle level (Goremykin et al., 2012; Wu et al., 2015; Wen et al., 2024). Nonetheless, there is an absence of complete mitochondrial genome data for this species needed for

phylogenetic studies. In this study, based on 18 mitochondrial common genes, *A. speltooides* and *A. thaliana* were selected as outgroups to construct a phylogenetic tree of 34 legume species, and the phylogenetic status of *M. ruthenica* was clarified. In addition, Phylogenetic analysis further revealed that *M. ruthenica* and *M. truncatula*, *M. polymorpha*, *M. sativa* ssp. *falcata*, *M. sativa* ssp. *sativa*, *M. arabica* and *M. lupulina* were clustered into the same branch in the phylogenetic tree, which further supported the taxonomic placement of *M. ruthenica* within *Medicago*. *M. ruthenica* was placed within the *Medicago* clade, but formed an independent branch. Phylogenetic analysis strongly supports a shared ancestry with other *Medicago* species, indicating *M. ruthenica* represents a distinct species within the genus (Xu et al., 2021). The branch lengths in phylogenetic trees reflect evolutionary divergence time. Here, *M. ruthenica* exhibited significantly longer branch lengths than did the other six congeneric species, suggesting either prolonged evolutionary divergence or later speciation compared with that of its relatives. The evolutionary history of *Medicago* likely involves survival challenges due to severe environmental changes. Chloroplast genome analysis of 61 *M. ruthenica* accessions revealed two major clades that diverged during the Last Glacial Period (Yan et al., 2021). Organelle-based molecular analyses can effectively explore genetic relationships and evolutionary processes in *M. ruthenica*, despite the limitations of high substitution rates and potential phylogenetic signal biases (Li et al., 2013).

This study expands the mitochondrial genome resources and provides novel insights into *M. ruthenica* classification. In conclusion, the results of this study not only reinforce certain theoretical foundations of traditional taxonomy, but also highlight the important value of mitochondrial genome in plant phylogenetic research. In addition, both maximum likelihood and Bayesian inference analyses were employed in this study, and the results showed high node support values, which indicated that mitochondrial genes had high reliability in analyzing the evolutionary relationship of legume species. At the same time, the topological structure of the phylogenetic tree constructed herein is consistent with previously reported legume phylogenies, which further verifies the credibility of the research results. Although the phylogenetic tree constructed based on mitochondrial genes in this study provides an important reference for the evolutionary relationship of legume species, there are still some limitations. At present, reports on the mitochondrial genome of legume plants are relatively scarce. Therefore, future studies should therefore aim to expand the sample size, which will help to more comprehensively and deeply analyze the evolutionary history of legume species and their complex evolutionary mechanisms.

5 Conclusions

The *M. ruthenica* mitochondrial genome spans 354,988 bp, approximately 2.80-fold larger than its chloroplast counterpart (126,939 bp), with a GC content of 45.13%. It contains 34 PCGs

18 tRNA genes, and 3 rRNA genes (*rrn18*, *rrn26*, *rrn5*). The mitogenome harbors 72 SSRs, with the dinucleotide (24) and tetranucleotide (21) repeats being the most abundant. The RSCU value of the codon UUA encoding Leu was the largest (1.55), the number of codon UUU encoding Phe was the largest (257–341), and the number of codon UAG encoding Ter was the least (4–5). Ka/Ks value analysis showed that genes such as *atp8*, *atp4*, *rps4*, *nad4*, *ccmB*, *ccmFc* and *nad9* showed significant positive selection in multiple legumes (Ka/Ks > 1). Furthermore, phylogenetic tree revealed *M. ruthenica* belongs to the genus *Medicago* and is a separate branch. These results provide a valuable model for studying *M. ruthenica* evolution and offer critical reference for mitochondria-associated traits in *M. ruthenica* breeding.

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

YT: Writing – original draft, Writing – review & editing. ZW: Methodology, Writing – review & editing. CT: Investigation, Writing – review & editing. YY: Formal analysis, Writing – review & editing. ZL: Writing – review & editing, Funding acquisition.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. The research was funded by Science & Technology Fundamental Resources Investigation Program (2024FY100302), Inner Mongolia Science and Technology Plan (2022YFHH0140), Hohhot Science and Technology Plan (2022-she-zhong-1-2).

Conflict of interest

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

The handling editor ZW declared a shared parent affiliation with the author(s) YT, ZW, CT, YY at the time of review.

Generative AI statement

The author(s) declare that no Generative AI was 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.2025.1648505/full#supplementary-material>

References

- AI, Y., Chen, J. B., Chen, Y. L., Wang, X. Q., Wang, N., Liang, W. Y., 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
- Arseneau, J., Steeves, R., and Laflamme, M. (2017). Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. *Mol. Ecol. Resour.* 17, 686–693. doi: 10.1111/1755-0998.12616
- Begum, Y., and Mondal, S. K. (2021). Comprehensive study of the genes involved in chlorophyll synthesis and degradation pathways in some monocot and dicot plant species. *J. Biomol. Structure. Dynamics.* 39, 2387–2414. doi: 10.1080/07391102.2020.1748717
- Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Benson, G. (1999). Tandem Repeats Finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573
- Berry, S. (2003). Endosymbiosis and the design of eukaryotic electron transport. *Biochim. Biophys. Acta (BBA) - Bioenerget.* 1606, 57–72. doi: 10.1016/S0005-2728(03)00084-7
- Campbell, T. A., and Baughan, G. R. (2002). Organelle based molecular analyses of the genetic relatedness of cultivated alfalfa (*Medicago sativa* L.) to *Medicago edgeworthii* sirjaev, and *Medicago ruthenica* (L.) Ledebour. *Euphytica* 125, 51–58. doi: 10.1023/A:1015796326675
- Čechová, J., Lýsek, J., Bartas, M., and Brázda, V. (2018). Complex analyses of inverted repeats in mitochondrial genomes revealed their importance and variability. *Bioinformatics* 34, 1081–1085. doi: 10.1093/bioinformatics/btx729
- Chen, H., and Skylaris, C. K. (2021). Analysis of DNA interactions and GC content with energy decomposition in large-scale quantum mechanical calculations. *Phys. Chem. Chem. Phys.* 23, 8891–8899. doi: 10.1039/D0CP06630C
- 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, H., Zeng, Y., and Yang, Y. (2020). Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nat. Commun.* 11, 2494. doi: 10.1038/s41467-020-16338-x
- Choi, I., Jansen, R., and Ruhlman, T. (2020). Caught in the Act: Variation in plastid genome inverted repeat expansion within and between populations of *Medicago minima*. *Ecol. Evol.* 10, 12129–12137. doi: 10.1002/ece3.6839
- Christensen, A. C. (2013). Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol. Evol.* 5, 1079–1086. doi: 10.1093/gbe/evt069
- Christensen, A. C. (2014). Genes and junk in plant mitochondria-repair mechanisms and selection. *Genome Biol. Evol.* 6, 1448–1453. doi: 10.1093/gbe/evu115
- 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
- Dyall, S. D., Brown, M. T., and Johnson, P. J. (2004). Ancient invasions: from endosymbionts to organelles. *Science* 304, 253–257. doi: 10.1126/science.1094884
- Fennoy, S. L., and Bailey-serres, J. (1993). Synonymous codon usage in *Zea mays* L. nuclear genes is varied by levels of C and G-Ending codons. *Nucl. Acids Res.* 21, 5294–5300. doi: 10.1093/nar/21.23.5294
- Figueroa, P., Gómez, I., Holuigue, L., Araya, A., and Jordana, X. (1999). Transfer of *rps14* from the mitochondrion to the nucleus in maize implied integration within a gene encoding the iron-sulphur subunit of succinate dehydrogenase and expression by alternative splicing. *Plant J.* 18, 601–609. doi: 10.1046/j.1365-3113.1999.00485.x
- Fontecilla-Camps, J. C. (2023). Reflections on the origin and early evolution of the genetic code. *ChemBioChem* 24, e202300048. doi: 10.1002/cbic.202300048
- Giegé, P., and Brennicke, A. (1999). RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl. Acad. Sci. U. S. A.* 96, 15324–15329. doi: 10.1073/pnas.96.26.15324
- Goremykin, V. V., Lockhart, P. J., Viola, R., and Velasco, R. (2012). The mitochondrial genome of *Malus Domestica* and the import-driven hypothesis of mitochondrial genome expansion in seed plants. *Plant J.* 71, 615–626. doi: 10.1111/j.1365-3113.2012.05014.x
- Gossau, A., Ruoff, P., Mohsenzadeh, S., Hobohm, U., and Rensing, L. (2001). Heat shock and oxidative stress-induced exposure of hydrophobic protein domains as common signal in the induction of *hsp68*. *J. Biol. Chem.* 276, 1814–1821. doi: 10.1074/jbc.M008280200
- Greiner, S., Lehwark, P., and Bock, R. (2019). Organellar genome DRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47, 59–64. doi: 10.1093/nar/gkz238
- Gualberto, J. M., Bonnard, G., Lamattina, L., and Grienemberger, J. M. (1991). Expression of the wheat mitochondrial *nad3-rps12* transcription unit: correlation between editing and mRNA maturation. *Plant Cell.* 3, 1109–1120. doi: 10.1105/tpc.3.10.1109
- Guisinger, M. M., Chumley, T. W., Kuehl, J. V., Boore, J. L., and Jansen, R. K. (2010). Implications of the plastid genome sequence of *Typha* (Typhaceae, Poales) for understanding genome evolution in *Poaceae*. *J. Mol. Evol.* 70, 149–166. doi: 10.1007/s00239-009-9317-3
- He, Y., Liu, W., and Wang, J. (2023). Assembly and comparative analysis of the complete mitochondrial genome of *Trigonella Foenum*-Graecum L. *BMC Genomics* 24, 756. doi: 10.1186/s12864-023-09865-6
- Hu, Y., Wu, L., Zhang, P., Wang, Z., Shang, J., and Huang, Y. (2021). Global view of dynamic expression and precise mapping of mitochondrial tRNAs-derived fragments during stressed conditions in *S. Pombe*. *Mitochondrion* 60, 219–227. doi: 10.1016/j.mito.2021.08.012
- 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
- Jia, W., and Higgs, P. G. (2008). Codon usage in mitochondrial genomes: distinguishing context-dependent mutation from translational selection. *Mol. Biol. Evol.* 25, 339–351. doi: 10.1093/molbev/msm259
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., Claude, W. D., Yi, T. S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 241. doi: 10.1186/s13059-020-02154-5
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von, H. A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Kanazawa, A., Tozuka, A., Kato, S., Mikami, T., Abe, J., and Shimamoto, Y. (1998). Small interspersed sequences that serve as recombination sites at the *cox2* and *atp6* loci in the mitochondrial genome of soybean are widely distributed in higher plants. *Curr. Genet.* 33, 188–198. doi: 10.1007/s002940050326
- Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108
- Keller, J., Rousseau-Gueutin, M., Martin, G. E., Morice, J., Boutte, J., Coissac, E., et al. (2017). The evolutionary fate of the chloroplast and nuclear *rps16* genes as revealed through the sequencing and comparative analyses of four novel legume chloroplast genomes from *Lupinus*. *DNA Res.* 24, 343–358. doi: 10.1093/dnares/dsx006
- Kempken, F. (2011). *Plant mitochondria* (New York Springer: Springer), 01, 03–19. doi: 10.1007/978-0-387-89781-3

- Kiktev, D. A., Sheng, Z., Lobachev, K. S., and Petes, T. D. (2018). GC content elevates mutation and recombination rates in the yeast *Saccharomyces cerevisiae*. *Proc. Nat. Acad. Sci. U.S.A.* 115, 7109–7118. doi: 10.1073/pnas.1807334115
- Knoop, V. (2013). Plant mitochondrial genome peculiarities evolving in the earliest vascular plant *Lineages*. *J. Systematics. Evol.* 51, 1–12. doi: 10.1111/j.1759-6831.2012.00228.x
- Konopiński, M. K. (2022). Average weighted nucleotide diversity is more precise than p_{xy} in estimating the true value of π from sequence sets containing missing data. *Mol. Ecol. Resour.* 23, 348–354. doi: 10.1111/1755-0998.13707
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Philippy, 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
- Kubo, T., and Newton, K. J. (2008). Angiosperm mitochondrial genomes and mutations. *Mitochondrion* 8, 5–14. doi: 10.1016/j.mito.2007.10.006
- Kumar, S., Kumari, S., and Shanker, A. (2020). In silico mining of simple sequence repeats in mitochondrial genomes of genus *Orthotrichum*. *JSR* 64, 179–182. doi: 10.37398/JSR.2020.640225
- Kurtz, S. (2001). REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633
- Kuzmin, E. V., Duvick, D. N., and Newton, K. J. (2005). A mitochondrial mutator system in maize. *Plant Physiol.* 137, 779–789. doi: 10.1104/pp.104.053611
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Latch, G. C. M. (1976). A revision of *melissitus* (Leguminosae). *Kew Bull.* 31, 495–502. doi: 10.1007/s12225-021-09991-y
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics* 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. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, H., Li, Z., Cai, L., Shi, W., Mi, F., and Shi, F. (2013). Analysis of genetic diversity of *ruthenia medic* (*Medicago ruthenica* (L.) Trautv.) in Inner Mongolia using ISSR and SSR Markers. *Genet. Resour. Crop Evol.* 60, 1687–1694. doi: 10.1007/s10722-012-9950-3
- Li, J., Li, H., Chi, E., Huang, F., Liu, L., Ding, Z., et al. (2018). Development of simple sequence repeat (SSR) markers in *Medicago ruthenica* and their application for evaluating out crossing fertility under open-pollination conditions. *Mol. Breeding* 38, 143. doi: 10.1007/s11032-018-0905-z
- Li, H. Y., Li, Z. Y., Shi, W. G., Li, X. Y., and Liu, L. (2010). Leaf anatomic indexes and the relations with drought resistance of the six forage grass of *Leguminosae*. *Acta Botan. Boreali-Occiden. Sin.* 30, 1989–1994.
- 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
- Linnaeus, C. (2015). Species plantarum: exhibentes plantas rite cognitae ad genera relatas. *Arch. Anim. Nutr.* 69, 442–454. doi: 10.1080/1745039X.2015.1093872
- Liu, L., Long, Q., Lv, W., Qian, J., Egan, A. N., Shi, Y., et al. (2024). Long repeat sequences mediated multiple mitogenome conformations of mulberries (*Morus* spp.), an important economic plant in China. *Genomics Commun.* 1, e005. doi: 10.48130/gcomm-0024-0005
- Liu, H., Yu, J., Yu, X., Zhang, D., Chang, H., Li, W., et al. (2021). Structural variation of mitochondrial genomes sheds light on evolutionary history of soybeans. *Plant J.* 108, 1456–1472. doi: 10.1111/tpj.15522
- Lowe, T. M., and Eddy, S. R. (1997). tRNAscan-SE: A Program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964. doi: 10.1093/nar/25.5.955
- Massa, A., Sobolev, V., Faustinelli, P., Talluri, S., Stalker, H., Lamb, M., et al. (2024). Genetic diversity, disease resistance and environmental adaptation of *Arachis duranensis* L.: new insights from landscape genomics. *PLoS One* 19, e0299992. doi: 10.1371/journal.pone.0299992
- Matsui, A., Nguyen, A., Nakaminami, K., and Seki, M. (2013). *Arabidopsis* non-coding RNA regulation in abiotic stress responses. *IJMS* 14, 22642–22654. doi: 10.3390/ijms141122642
- Melton, J. T., Leliaert, F., Tronholm, A., and Lopez-Bautista, J. M. (2015). The complete chloroplast and mitochondrial genomes of the green macroalga *Ulva* Sp. UNA00071828 (Ulvophyceae, Chlorophyta). *PLoS One* 10, e0121020. doi: 10.1371/journal.pone.0121020
- Mousavi-Derazmahalleh, M., Bayer, P., Hane, J., Valliyodan, B., Nguyen, H., Nelson, M., et al. (2018). Adapting legume crops to climate change using genomic approaches. *Plant Cell Environ.* 42, 6–19. doi: 10.1111/pce.13203
- Muyle, A., Serres-Giardi, L., Ressayre, A., Escobar, J., and Glémin, S. (2011). GC-biased gene conversion and selection affect GC content in the *Oryza* genus (rice). *Mol. Biol. Evol.* 28, 2695–2706. doi: 10.1093/molbev/msr104
- Nishikawa, T., Vaughan, D. A., and Kadowaki, K. (2005). Phylogenetic analysis of *Oryza* species, based on simple sequence repeats and their flanking nucleotide sequences from the mitochondrial and chloroplast genomes. *Theor. Appl. Genet.* 110, 696–705. doi: 10.1007/s00122-004-1895-2
- Park, S., Gilmour, S. J., Grumet, R., and Thomashow, M. F. (2018). CBF-dependent and CBF-independent regulatory pathways contribute to the differences in freezing tolerance and cold-regulated gene expression of two arabidopsis ecotypes locally adapted to sites in Sweden and Italy. *PLoS One* 13, e0207723. doi: 10.1371/journal.pone.0207723
- Parvathy, S. T., Udayasuriyan, V., and Bhadana, V. (2022). Codon usage bias. *Mol. Biol. Rep.* 49, 539–565. doi: 10.1007/s11033-021-06749-4
- Peden, J. F. (2000). *Analysis of codon usage. Doctoral dissertation* (England: University of Nottingham).
- 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. doi: 10.1186/s13007-019-0435-7
- Rehman, O., Uzair, M., Chao, H., Khan, M. R., and Chen, M. (2022). Decoding RNA editing sites through transcriptome analysis in rice under alkaline stress. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.892729
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029
- 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
- Ruprecht, C., Lohaus, R., Vanneste, K., Mutwil, M., Nikoloski, Z., Van, D. P. Y., et al. (2017). Revisiting ancestral polyploidy in plants. *Sci. Adv.* 3, e1603195. doi: 10.1126/sciadv.1603195
- Sablok, G., Yang, K., Chen, R., and Wen, X. (2017). tRNA derived smallRNAs: smallRNAs repertoire has yet to be decoded in Plants. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01167
- Semerikov, V. L., Semerikova, S. A., Khrunyk, Y. Y., and Putintseva, Y. A. (2022). Sequence capture of mitochondrial genome with PCR-generated baits provides new insights into the biogeography of the genus *abies* mill. *Plants* 11, 762. doi: 10.3390/plants11060762
- Shakyawar, S. K., Joshi, B. K., and Kumar, D. (2009). SSR repeat dynamics in mitochondrial genomes of five domestic animal species. *Bioinformation* 4, 158–163. doi: 10.6026/97320630004158
- Sharma, N., Pasala, M. S., and Prakash, A. (2019). Mitochondrial DNA: epigenetics and environment. *Environ. Mol. Mutagen.* 60, 668–682. doi: 10.1002/em.22319
- Shu, Y., Li, W., Zhao, J., Liu, Y., and Guo, C. (2018). Transcriptome sequencing and expression profiling of genes involved in the response to abiotic stress in *Medicago ruthenica*. *Genet. Mol. Biol.* 41, 638–648. doi: 10.1590/1678-4685-gmb-2017-0284
- Sloan, D. B., Alverson, A. J., Chukalovcak, J. P., Wu, M., McCauley, D. E., Palmer, J. D., et al. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* 10, e1001241. doi: 10.1371/journal.pbio.1001241
- Small, E., and Jomphé, M. (1989). A synopsis of the genus *Medicago* (Leguminosae). *Can. J. Bot.* 67, 3260–3294. doi: 10.1139/b89-405
- Smith, D. R., and Keeling, P. J. (2015). Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10177–10184. doi: 10.1073/pnas.1422049112
- Soják, S. (1982). *Melioides*, a new genus of the tribe *Trigonelleae* (Leguminosae). *Preslia* 54, 115–122. doi: 10.1038/s41597-021-00997-6
- Stamatakis, A. (2006). RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatic* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446
- Takezawa, D., and Minami, A. (2004). Calmodulin-binding proteins in bryophytes: identification of abscisic acid-, cold-, and osmotic stress-induced genes encoding novel membrane-bound transporter-like proteins. *Biochem. Biophys. Res. Commun.* 317, 428–436. doi: 10.1016/j.bbrc.2004.03.052
- Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577. doi: 10.1080/10635150701472164
- Tian, C. Y., Wu, Z. N., Li, X. S., and Li, Z. Y. (2021). Codon usage bias of chloroplast genome in *M. Ruthenica*. *Acta Agrestia. Sinica.* 29 (12), 2678–2684. doi: 10.11733/j.issn.1007-0435.2021.12.006
- Unsel, M., Marienfeld, J. R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366, 924 nucleotides. *Nat. Genet.* 15, 57–61. doi: 10.1038/ng0197-57
- Wang, J., Kan, S., Liao, X., Zhou, J., Tembrock, L. R., Daniell, H., et al. (2024a). Plant organellar genomes: much done, much more to do. *Trends Plant Sci.* 29, 754–769. doi: 10.1016/j.tplants.2023.12.014
- Wang, Y., Sun, N., Shi, W. Q., Sun, L., Hao, M., Bi, C., et al. (2023). Assembly and comparative analysis of the complete mitochondrial genome of *ilex macrocarpa*. *Forests* 14, 2372. doi: 10.3390/f14122372
- Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCS-X: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49–e49. doi: 10.1093/nar/gkr1293
- 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. *Genom. Proteomics Bioinf.* 8, 77–80. doi: 10.1016/S1672-0229(10)60008-3

- Wang, Z. F., Zhang, Y., Zhong, X. J., Kang, N., and Cao, H. L. (2021). The complete mitochondrial genome of *Ormosia boluoensis*. *Mitochondrial. DNA Part B*. 6, 2109–2111. doi: 10.1080/23802359.2021.1920503
- Wang, S., Qiu, J., Sun, N., Han, F., Wang, Z., et al. (2025). Characterization and comparative analysis of the first mitochondrial genome of *Michelia* (Magnoliaceae). *Genomics Communications* 2, e001. doi: 10.48130/gcomm-0025-0001
- Wang, J., Zou, Y., Mower, J. P., Wayne, R., and Wu, Z. Q. (2024b). Rethinking the mutation hypotheses of plant organellar DNA. *Genomics Commun.* 1, e003. doi: 10.48130/gcomm-0024-0003
- Wen, J., Guo, M. W., Zhu, L., Li, Z. Y., Li, H. Y., Wei, P., et al. (2024). Analysis of chloroplast genome and phylogeny of yellow flowered type *M. Ruthenica*. *Chin. J. Grassland*. 46, 10–18. doi: 10.16742/j.zgdx.20230367
- 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
- Wickham, H. (2009). *Ggplot2: elegant graphics for data analysis* (New York: Springer).
- Wu, Z., Cuthbert, J. M., Taylor, D. R., and Sloan, D. B. (2015). The massive mitochondrial genome of the angiosperm *Silene Noctiflora* is evolving by gain or loss of entire chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10185–10191. doi: 10.1073/pnas.1421397112
- 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
- Wu, Z., Oeck, S., West, A. P., Mangalhar, K. C., Sainz, A. G., Newman, L. E., et al. (2019). Mitochondrial DNA stress signalling protects the nuclear genome. *Nat. Metab.* 1, 1209–1218. doi: 10.1038/s42255-019-0150-8
- Wu, Z. Y., Raven, P. H., and Hong, D. Y. (2010). *Flora of China* Vol. 10 (Beijing: Science press. Missouri Botanical Garden Press (St. Louis), 554–555.
- Wu, Z., Yang, T., Qin, R., and Liu, H. (2023). Complete mitogenome and phylogenetic analysis of the *Carthamus Tinctorius* L. *Genes* 14, 979. doi: 10.3390/genes14050979
- Wu, C., Zinshteyn, B., Wehner, K., and Green, R. (2019). High-resolution ribosome profiling defines discrete ribosome elongation states and translational regulation during cellular stress. *Mol. Cell*. 735, 959–970. doi: 10.1016/j.molcel.2018.12.009
- Wuriya, Y., Yong, S. P., and Bao, G. P. (1994). A comparative study on ecobiological characteristics of *Melilotoides ruthenica*. *Grassland. China*. 2, 1–7. doi: 10.16742/j.zgdx.20230367
- Wynn, E. L., and Christensen, A. C. (2019). Repeats of unusual size in plant mitochondrial genomes: identification, incidence and evolution. *G3 Genes Genomes Genet.* 9, 549–559. doi: 10.1534/g3.118.200948
- Xie, J., Mao, J., Li, Z., Liang, C., Ou, W., Tang, J., et al. (2021). Complete chloroplast genome of a high-quality forage in north China, *Medicago ruthenica* (Fabaceae: trifolieae). *Mitochondrial. DNA Part B*. 6, 29–30. doi: 10.1080/23802359.2020.1845578
- Xiong, Y., Yu, Q., Xiong, Y., Zhao, J., Lei, X., Liu, L., et al. (2022). The complete mitogenome of *Elymus sibiricus* and insights into its evolutionary pattern based on simple repeat sequences of seed plant mitogenomes. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.802321
- Xu, H., Li, Y., Shang, X., Zhong, H., Dong, K., and Xia, F. (2021). The complete chloroplast genome of *Medicago ruthenica* cv. 'Taihang' (Fabaceae) from Shanxi, China. *Mitochondrial. DNA Part B*. 6, 2688–2690. doi: 10.1080/23802359.2021.1966328
- Yan, W., Shi, W., Tian, Q., Li, Z., and Gao, H. (2021). The complete chloroplast genome sequence of *Melilotoides ruthenica*: structural comparative and phylogenetic analysis in Leguminosae. *IOP. Conf. Ser. Earth Environ. Sci.* 697, 12014. doi: 10.1088/1755-1315/697/1/012014
- Yang, Z., and Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* 15, 496–503. doi: 10.1016/S0169-5347(00)01994-7
- Yang, H., Li, W., Yu, X., Zhang, X., Zhang, Z., Liu, Y., et al. (2021). Insights into molecular structure, genome evolution and phylogenetic implication through mitochondrial genome sequence of *Gleditsia Sinensis*. *Sci. Rep.* 11, 14850. doi: 10.1038/s41598-021-93480-6
- Yang, M., Sui, J., and Li, X. (2023). The first complete mitogenome of *Acharax* sp. (protobranchia, solemyida, solemyidae): Comparisons with other solemyidae bivalves and deep-sea adaptive characteristics. *J. Oceanol. Limnol.* 41, 2374–2390. doi: 10.1007/s00343-022-2217-8
- Yang, Y., Wang, X., and Shi, Z. (2024). Comparative study on codon usage patterns across chloroplast genomes of eighteen *Taraxacum* species. *Horticulturae* 10, 492. doi: 10.3390/horticulturae10050492
- Yin, M., Zhang, S., Du, X., Mateo, R. G., Guo, W., Li, A., et al. (2021). Genomic analysis of *Medicago ruthenica* provides insights into its tolerance to abiotic stress and demographic history. *Mol. Ecol. Resour.* 21, 1641–1657. doi: 10.1111/1755-0998
- Yu, T., Sun, L., Cui, H., Liu, S., Men, J., Chen, S., et al. (2018). The complete mitochondrial genome of a tertiary relict evergreen woody plant *Ammopiptanthus Mongolicus*. *Mitochondrial. DNA Part B*. 3, 9–11. doi: 10.1080/23802359.2017.1413301
- Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., et al. (2020). PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355. doi: 10.1111/1755-0998.13096
- Zhang, S., Wang, J., He, W., Kan, S., Liao, X., Jordan, D. R., et al. (2023). Variation in mitogenome structural conformation in wild and cultivated lineages of sorghum corresponds with domestication history and plastome evolution. *BMC Plant Biol.* 23, 91. doi: 10.1186/s12870-023-04104-2
- Zhao, C. X., Zhu, R. L., and Liu, Y. (2016). Simple sequence repeats in *Bryophyte* mitochondrial genomes. *Mitochondrial. DNA Part A*. 27, 191–197. doi: 10.3109/19401736.2014.880889
- Zwonitzer, K. D., Tressel, L. G., Wu, Z., Kan, S., Broz, A. K., Mower, J. P., et al. (2024). Genome copy number predicts extreme evolutionary rate variation in plant mitochondrial DNA. *Proc. Natl. Acad. Sci. U. S. A.* 121, e2317240121. doi: 10.1073/pnas.2317240121