BMC Plant Biology

RESEARCH

Open Access

Comprehensive analyses of different putative ploidy levels in organelle genomes of an important medicinal plant *Polygonatum kingianum* Collett & Hemsl.

Jing Zhao[3†], Zi-Han Chen[1†], Ming-Xian Zhang[1], Li-Hua Wang[1], Xiao-Yi Yang[1], Jia-Guan Wang[4], You-Yong Zhu[1], Xing-Yao Li[2*] and Yu Zhao[1*]

Abstract

Background *Polygonatum kingianum*, a key species in traditional Chinese medicine, is increasingly valued for its medicinal and nutritional properties. However, wild resources are declining due to over-exploitation, necessitating genomic studies for conservation.

Results Flow cytometry analysis based on genome size revealed the presence of both putative diploid and tetraploid *P. kingianum*. By further integrating sequencing data from both Illumina and Oxford Nanopore Technologies (ONT), the organelle genomes were assembled. While the chloroplast genomes of different putative ploidy levels of *P. kingianum* were highly conserved in structure and features, the mitogenome of putative tetraploid *P. kingianum* consisted of two chromosomes, whereas that of putative diploid *P. kingianum* contained only one. Mitochondrial gene annotation showed that putative tetraploid *P. kingianum* possesses one additional copy each of the *nad*3 and *nad*5 genes compared to the putative diploid *P. kingianum*. Analyses of repetitive sequences indicated that, although no obvious correlation was observed between tandem repeats or dispersed repeats and different putative ploidy levels in the chloroplast genomes or between tandem repeats and different putative ploidy levels in the mitogenomes, simple sequence repeats in the organelle genomes correlated with ploidy variation of *P. kingianum*. Investigation of RNA editing sites revealed ploidy-dependent variations in the *cox*2 and *nad*3 genes of the mitogenomes across different ploidy levels of *P. kingianum*, suggesting that differences in editing sites may be linked to the adaptability of the different ploidies. Phylogenetic analyses based on mitochondrial genes and nucleotide substitution rate analysis showed no significant differences or distinct variations among different ploidy levels of *P. kingianum*, supporting that the morphologically highly variable *P. kingianum* comprises only a single species.

†Jing Zhao and Zi-Han Chen contributed to the work equally and should be regarded as co-first authors.

*Correspondence:
Xing-Yao Li
13529802255@163.com
Yu Zhao
zhaoyu@ynau.edu.cn

Full list of author information is available at the end of the article

Conclusions In this study, we successfully assembled the organelle genomes of *P. kingianum* with different putative ploidy levels and conducted comparative genomic analyses. We found that the mitogenomes of *P. kingianum* are structurally complex and variable, playing crucial roles in evolution and adaptation. This study not only enhances our understanding of the organelle genome characteristics of different putative ploidy levels within the same species, but also lays the foundation for the subsequent development and utilization of the medicinal value of *Polygonatum*.

Keywords Asparagaceae, Organelle genome, Phylogeny, SSRs, Putative ploidy

Introduction

The genus *Polygonatum* Mill. is a medium-sized lineage within the Asparagaceae family, comprising approximately 65 species primarily distributed across the Northern Hemisphere [1–4]. In China, *Polygonatum* contains about 39 species, 20 of which are endemic [1]. Most *Polygonatum* species are used in traditional Chinese medicine (TCM) and possess significant medicinal and edible value [1, 2]. *Polygonatum kingianum* Collett & Hemsl., a representative geo-authentic herb, has garnered considerable attention in both the TCM and functional food industries [5, 6]. The *Chinese Pharmacopoeia* lists *P. kingianum* as one of the official species for the medicinal material "Huangjing" [6]. The rhizomes of *P. kingianum* are rich in bioactive compounds, including polysaccharides, saponins, and amino acids, which confer pharmacological properties such as hypoglycemic effects, yin-nourishing, kidney-tonifying, and lung-moistening activities [5, 7, 8]. With growing consumer demand for natural therapeutics, the cultivation and commercial value of *P. kingianum* have expanded rapidly. However, wild populations face severe depletion due to overharvesting and habitat degradation [1, 5, 6], highlighting the need for genomic studies to facilitate conservation and sustainable utilization.

In green plants, mitochondria and chloroplast are two organelles, in addition to the nucleus, that contain genetic information [9–11]. Unlike the biparentally inherited nuclear genome, the organelle genomes of most plants are maternally inherited and are not influenced by hybridization, playing a crucial role in studying plant origin and evolution [12, 13]. Recently, phylogenetic analyses based on the mitogenomes of *Ceratophyllum* and *Chloranthus* have enhanced our understanding of the angiosperm tree of life [14]. However, significant debate remains over whether *P. kingianum* represents a single species or multiple species [1, 15, 16]. Traditionally, *P. kingianum* is not only a species with extensive morphological variation but also is a polyploid complex with different chromosome numbers (2n = 26, 30, 32, 64) [1]. Wang et al. [16] analyzed the complete chloroplast genomes of *P. kingianum* with different flower colors and revealed that the species comprises three distinct clades. Although amount of chloroplast genomes of TCM have been sequenced, research on mitogenomes has significantly lagged [17]. This is primarily due to the characteristics of mitogenomes, such as extensive repetitive sequences, lower substitution rates, variation in sizes, intracellular gene transfer, and abundant RNA editing sites [18–21]. In recent years, with the advancement of third-generation sequencing technologies (e.g., PacBio HiFi and Oxford Nanopore), research on plant mitogenomes has progressed steadily, laying an important foundation for understanding plant evolution, species identification, and genetic diversity [13, 14, 22].

Compared to the highly conserved quadripartite structure found in the chloroplast genomes of most plants, the structure of the mitogenomes exist in branched linear or circular forms [23, 24]. This structural diversity in mitogenomes offers a distinct perspective for deepening our understanding of plant evolutionary history and phylogenetic relationships [13, 14]. Previous studies have shown that chloroplast genomes within a single species, or even within a family, typically exhibit only minor size variations [9]. However, significant differences in mitogenome sizes, and even the presence of distinct structural conformations, can exist among different individuals of the same species [25]. Despite a growing number of mitogenomes being sequenced and assembled, a substantial gap remains compared to chloroplast genomes, with the mitogenomes of many species still uncharacterized [26]. This gap severely hinders our understanding of structural variation in plant mitogenomes.

With these questions in minds, the aims of this study were to: (1) assemble the different putative ploidy levels complete mitogenome of *P. kingianum* using an integrated sequencing approach combining Illumina and Nanopore technologies; (2) characterize the organelle genome architecture, gene content, repetitive elements, codon usage patterns, substitution rates, and RNA editing events; and (3) clarify the relationships of *P. kingianum* across different putative ploidy levels.

Materials and methods

Plant materials and flow cytometry

Fresh and healthy leaves from six individuals of *P. kingianum* were collected from the greenhouse of Yunnan Agricultural University, Yunnan Province, China. The samples were immediately frozen in liquid nitrogen and then stored at − 80 °C. The voucher specimens were deposited at the Herbarium of Yunnan Agricultural University. Ploidy levels and absolute genome sizes of *P. kingianum*

were estimated using flow cytometry following the protocol of Doležel et al. [27]. *Zea mays* (genome size = ca. 2.3 Gb) was used as internal standard. DNA quantities were measured using a BD FACScalibur laser flow cytometer with Modifit v3.0. DNA ploidy and absolute genome sizes were determined based on the basis of the sample/standard ratio.

DNA extraction and sequencing

Two individuals with clear and stable ploidy signals were selected for sequencing (accession numbers Y. Zhao et al. PK8 and Y. Zhao et al. PK11). Total genomic DNA was extracted using an improved Cetyl Trimethyl Ammonium Bromide (CTAB) method [28]. The DNA samples were sent to Beijing Novogene Technology Co., Ltd. (Tianjin, China) for library construction and sequencing using both Illumina short-read and Oxford Nanopore long-read platforms. For Illumina short-read sequencing, a paired-end library with an insert size of 350 bp was constructed, and sequencing was performed on the Illumina NovaSeq X Plus platform. The raw Illumina data generated twenty gigabases (20 Gb). The raw reads were subsequently quality-trimmed using Fastp v0.23.1 with default parameters [29]. For Oxford Nanopore long-read sequencing, the library was sequenced on a PromethION platform. Raw data were processed using NanoPack (https://github.com/wdecoster/nanopack) to generate 25 Gb of data (mean read length 22,375 bp; voucher: Y. Zhao et al. PK11) and 42 Gb of data (mean read length 20,699 bp; voucher: Y. Zhao et al. PK8).

Organelle genomes assembly and annotation

For chloroplast genome assembly, GetOrganelle v1.7.7.1 [30] was used to assemble the Illumina paired-end clean reads with default parameters. Chloroplast genomes were annotated using CPGAVAS2 [31] and manually adjusted the start and stop codons and the exon–intron boundaries of genes in Geneious Prime 2021.2.2 referring to published chloroplast genome of *P. kingianum* (PQ611457). For mitogenome assembly, a hybrid strategy was employed initially, GetOrganelle was used to assemble mitochondrial reads from the short reads with parameters "-R 80 -k 21,45,65,85,105,115 -P 1000000 -F embplant_mt" and the complete mitogenome of *Chlorophytum comosum* (MW411187) as a reference. Subsequently, both the gfa file (generated by GetOrganelle) and/or the filtered ONT reads were used as input files for de novo assembly of mitogenome long reads using GSAT v1.1.2 [32] and PMAT v1.5.3 [33], respectively. Bandage v0.8.1 [34] was employed to visualize and disentangle the raw assembly graph of mitogenomes. Mitogenomes were annotated using the IPMGA webserver (www.1kmpg.cn/mga), and all annotations were manually checked and corrected for start and stop codons of protein-coding genes using Geneious Prime. tRNA genes of organelle genomes were identified with tRNAscan-SE v2.0 [35]. Circular genome maps were drawn with the OrganellarGenomeDRAW tool [36].

Comparative genomic analyses

Firstly, we analyzed three different types (simple sequence repeats: SSRs, dispersed repeats: DRs, tandem repeats: TRs) of repeat sequences in both chloroplast genomes and mitogenomes. SSRs were identified using the MISA online tool [37], with motif sizes ranging from one to six nucleotides. The repeat unit thresholds were set to 8, 4, 4, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats, respectively. Four types of DRs, namely forward (F), palindromic (P), reverse (R), and complement (C) repeats, were searched by the online version of REPuter with the following parameters: minimal repeats size = 30, and hamming distance = 3 [38]. TRs were identified using Tandem Repeats Finder [39], with alignment parameters set to 2 (match) and 7 (mismatches/indels).

Secondly, the protein coding genes sequence of the *P. kingianum* mitogenomes were extracted using PhyloSuite v1.2.2 [40]. CodonW v1.4.2 (http://codonw.sourceforge.net/) was applied to analyze the relative synonymous codon usage (RSCU) of the protein coding genes. The results were plotted as a heatmap using OriginPro v2022b (OriginLab Corporation, Northampton, MA, USA). As well, the non-synonymous (Ka) and synonymous (Ks) substitution rates (Ka/Ks) of mitogenome were analyzed for shared protein coding genes between *P. kingianum* and *Chlorophytum comosum* through Ka/Ks Calculator v2.0 [41]. In addition, the extremely high frequency of RNA editing is often observed in the mitogenome of plant [21]. Predicted the RNA editing sites in the protein coding regions of the mitogenomes using the online plant RNA editing prediction and analysis tool PREPACT3 [42]. Predictions were conducted under BLASTX mode with default parameters, and the mitochondrial protein database of *Amborella trichopoda* (KF754799) was selected as reference.

Molecular phylogenetic analyses

Thirty-seven complete mitogenome sequences from seven orders (Alismatales, Arecales, Asparagales, Magnoliales, Pandanales, Poales, and Ranunculales) were downloaded from NCBI, representing 33 genera. Among these, two species (*Aconitum kusnezoffii* and *Pulsatilla dahurica*) of Ranunculales were selected as outgroups (Table S1). These sequences, together with the two newly sequenced *P. kingianum* mitogenomes, were used to construct a phylogenetic tree. Phylogenetic relationships were reconstructed using 41 protein coding genes extracted through PhyloSuite. Each gene was first aligned and then trimmed using the plugins Mafft v7.450 (set

E-INS-i) [43] and trimAI v1.3 (set -automated1) [44]. The aligned sequences were then concatenated using the "Concatenate Sequence" plugin in PhyloSuite. The nucleotide substitution rate variation models of Maximum Likelihood (ML) and Bayesian Inference (BI) were fitted using the ModelFinder tool under the corrected Akaike Information Criterion (AICc) in IQ-tree v2.1.3 [45]. ML analyses were performed with 5,000 ultrafast bootstraps [46] replicates in IQ-tree. BI was conducted using MrBayes v3.2.2 [47]. Four Markov chain Monte Carlo chains were run, each beginning with a random tree and sampling one tree every 100 generations of 2,000,000 generations, and the first 25% of samples were discarded as burn-in. The standard deviation of splits frequencies below 0.001, and the Markov Chain Monte Carlo (MCMC) output was examined to check for convergence and to ensure that all of the effective sample size (ESS) values were ≥ 200. Maximum Likelihood Bootstrap Support values (ML-BS) and Bayesian Inference Posterior Probabilities (BI-PP) ≥ 80/0.90 were defined as strongly supported.

Results

Genome sizes and cytotypes

The results obtained from flow cytometry revealed that six individuals of *P. kingianum* only were two cytotypes (Table S2). The DNA amount of *P. kingianum* was estimated to range from 5.18 Gb (voucher: Y Zhao et al. PK11) to 13.28 Gb (voucher: Y Zhao et al. PK10) by using *Zea mays* as an internal standard (Table S2). Although the flow cytometry results were not supplemented by conventional chromosome counts to verify the ploidy levels, we found a nearly twofold change in genome sizes. The genome sizes of PK8, PK9, PK10, and PK13 were putative recognized as a putative tetraploid, while the other of *P. kingianum* (voucher: Y Zhao et al. PK7 and Y Zhao et al. PK11) were putative recognized as diploid (Table S2).

Basic characteristics of *Polygonatum kingianum* organelle genomes

We successfully assembled the complete *P. kingianum* mitogenome based on both Illumina and ONT reads (Figs. 1a, b, c, f; S1). The mitogenomes of *P. kingianum* showed a complex assembly graph mediated by repeats (Fig. S1a, c). Through de-looping (Fig. S1b, d), the putative tetraploid *P. kingianum* was split into two chromosomes (Figs. 1a, b, f; S1b; Table 1), whereas the putative diploid *P. kingianum* contained only one chromosome (Figs. 1c, f; S1d; Table 1). Although both ploidy types of *P. kingianum* encode only 54 identical genes, including 18 tRNAs, 3 rRNAs, and 32 protein-coding genes, the copy numbers of these genes vary between ploidy levels (Table S3). Of the mitogenome of putative tetraploid *P. kingianum*, *atp*1 has three copies, while *nad*3, *nad*5, *rpl*16, *sdh*4, and *trnf*M-CAU each have two copies (Fig. 1a, b, f; Table S3). In contrast, in the mitogenome of putative diploid *P. kingianum* has three copies of *atp*1, while only *rpl*16, *sdh*4, and *trnf*M-CAU have two copies each (Fig. 1c, f; Table S3). Additionally, although the GC content of both mitogenomes is 46.30%, their sizes differ significantly (Fig. 1a, b, c, f; Table 1). The mitogenome size of putative tetraploid *P. kingianum* is only 643,467 bp, while that of putative diploid *P. kingianum* is up to 679,406 bp (Fig. 1a, b, c, f; Table 1).

Compared to the mitogenomes, the chloroplast genomes of *P. kingianum* with different putative ploidy levels are highly conserved, exhibiting a typical quadripartite structure that includes a large single-copy region (LSC), a small single-copy region (SSC), and a pair of inverted repeat regions (IRs) (Fig. 1d, e, g). Both chloroplast genomes of *P. kingianum* encode 131 genes, including 8 tRNAs, 38 rRNAs, and 85 protein coding genes (Fig. 1d, e, g; Table 1). Moreover, there are minimal differences in both GC content (37.70%) and chloroplast genome sizes (155,794–155,832 bp) (Fig. 1d, e, g; Table 1).

Repeat sequences analyses

Analyses of repetitive sequences revealed a strong correlation between the number of SSRs and the putative ploidy levels in *P. kingianum*. Specifically, the mitochondrial and chloroplast genomes of putative diploid *P. kingianum* contained 105 and 159 SSRs, respectively (Table S4). In contrast, the number of SSRs in putative tetraploid *P. kingianum* nearly doubled, with 226 in the mitogenome and 202 in the chloroplast genome (Table S4). Furthermore, the distribution of different SSR types in the mitogenome of *P. kingianum* also exhibited a clear ploidy-related pattern (Figs. 2a, c, S2a, c; Table S4). For example, the counts of monomers, dimers, trimers, and tetramers in the putative tetraploid were essentially double those in the putative diploid (Fig. 2a, c; Table S4). In the chloroplast genomes, the counts of different SSR types were significantly higher in the putative tetraploid *P. kingianum* than in putative diploid *P. kingianum*, but no clear ploidy correlation pattern was observed (Fig. S2a, c; Table S4). In addition, dimers were the most abundant in the mitogenomes of *P. kingianum* (Fig. 2a, c; Table S4), followed by monomers, tetramers, and trimers (Fig. 2a, c; Table S4). However, monomers were the most frequent within the chloroplast genomes of *P. kingianum*, followed by dimers, tetramers, and then trimers (Fig. S2a, c; Table S4).

TRs of 0–10 bp were found only in the chloroplast genomes of *P. kingianum* and were absent from the mitogenomes (Figs. 2b, S2b; Table S5). Conversely, TRs longer than 30 bp were exclusively distributed in the mitogenomes and not found in the chloroplast genomes

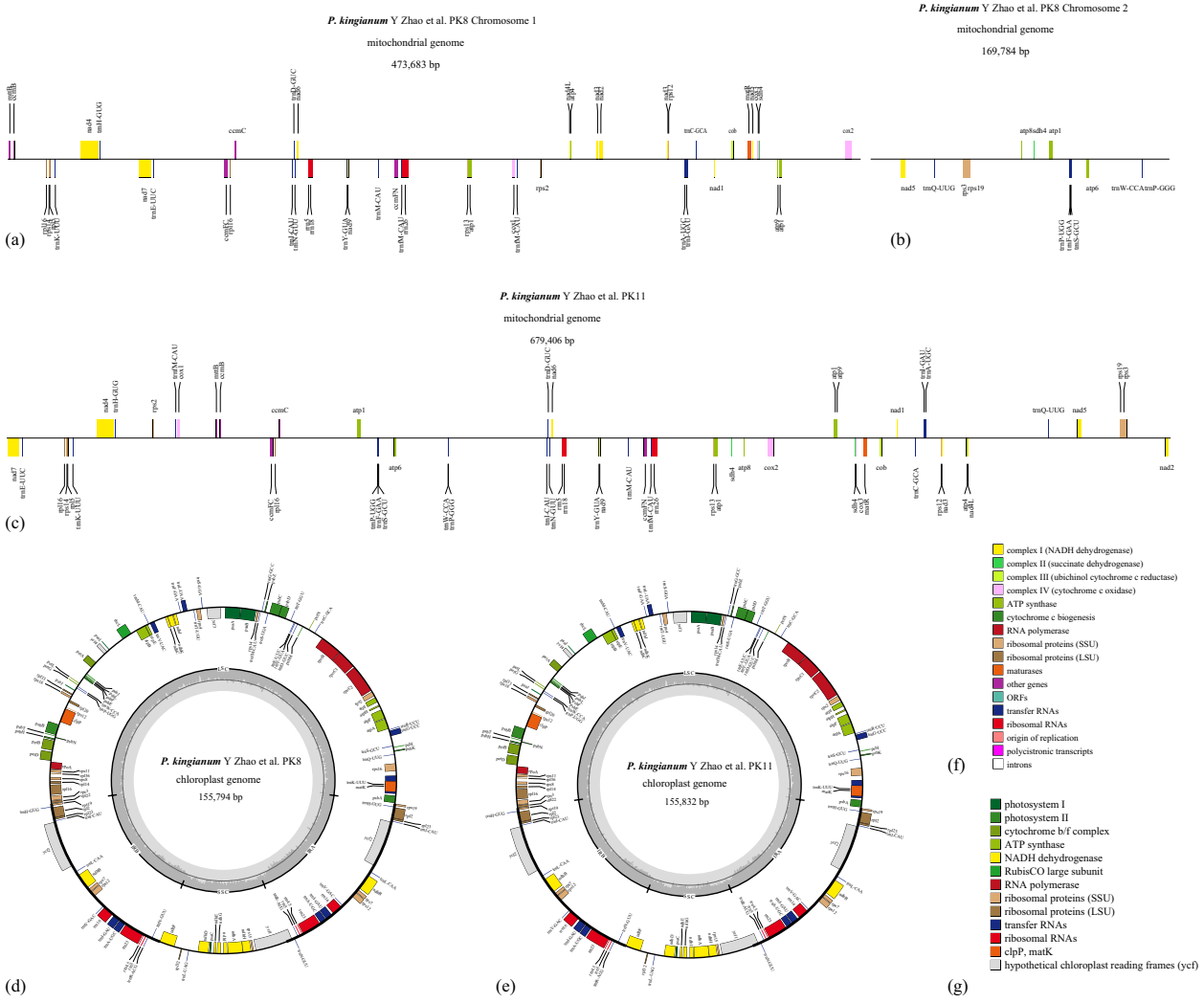

Fig. 1 The organelle genome maps. **a**, **b** The mitogenome map of putative tetraploid *P. kingianum*. **c** The mitogenome map of putative diploid *P. kingianum*. **d** The chloroplast genome map of putative tetraploid *P. kingianum*. **e** The chloroplast genome map of putative diploid *P. kingianum*. **f** The legend for mitochondrial genes belonging to different functional groups. **g** The legend for plastid genes belonging to different functional groups. Genes on the outside of the map are transcribed clockwise, and genes on the inside are transcribed counterclockwise

Table 1 Organelle genomes features and putative ploidy levels of *Polygonatum kingianum*

Species	*Polygonatum kingianum*	*Polygonatum kingianum*	*Polygonatum kingianum*	*Polygonatum kingianum*
Voucher	Y Zhao et al. PK8	Y Zhao et al. PK8	Y Zhao et al. PK11	Y Zhao et al. PK11
Organelle	chloroplast genome	mitogenome	chloroplast genome	mitogenome
Putative ploidy	tetraploid	tetraploid	diploid	diploid
Chromosome	1	2	1	1
Size (bp)	155,794	643,467	155,832	679,406
Total GC content (%)	37.70	46.30	37.70	46.30
Genes	131	61	131	59
tRNA	8	19	8	19
rRNA	38	3	38	3
Protein-coding genes	85	39	85	37

(Figs. 2b, S2b; Table S5). As well, TRs in the 51–60 bp range are present in putative tetraploid *P. kingianum* but absent in putative diploid *P. kingianum*, whereas repeats in the 61–70 bp range are present in putative diploid *P. kingianum* but absent in putative tetraploid *P. kingianum*. Although organelle genomes of putative diploid *P. kingianum* generally contained a higher number of TRs across various length categories compared to putative tetraploid

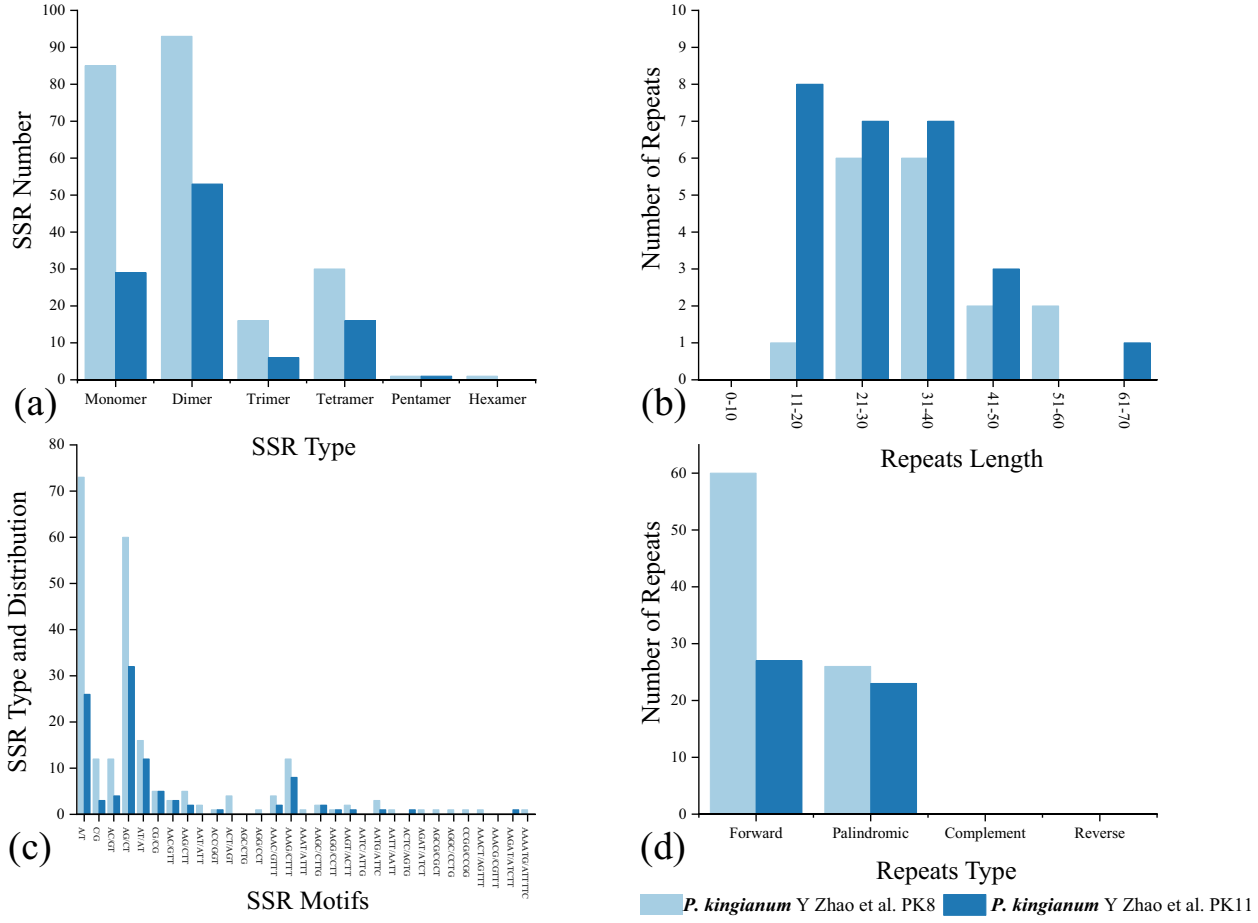

Fig. 2 Repeat sequences analyses of different putative ploidy levels of *Polygonatum kingianum* mitogenomes. **a** The frequency of identified simple sequence repeats. **b** Statistics of tandem repeats sequence with different lengths. **c** Type and quantity of simple sequence repeat motifs. **d** Statics of dispersed repeat sequence types

P. kingianum (Figs. 2b, S2b; Table S5), the total number of TRs in the chloroplast genomes was similar between ploidies (Fig. S2b; Table S5). In contrast, the mitogenome of putative diploid *P. kingianum* contained 26 TRs in total, while the putative tetraploid *P. kingianum* mitogenome had only 17 (Fig. 2b; Table S5). Regarding DRs, the complement repeats were absent in both organelle genomes (Figs. 2d, S2d; Table S6). Reverse repeats were present only in the chloroplast genomes and not in the mitogenomes (Figs. 2d, S2d; Table S6). Furthermore, forward repeats dominated in the mitogenomes, whereas palindromic repeats were predominanted in the chloroplast genomes (Figs. 2d, S2d; Table S6). Notably, the number of DRs in both the chloroplast genomes and mitogenomes were higher in putative tetraploid *P. kingianum* than in the putative diploid (Figs. 2d, S2d; Table S6).

Codon bias usage and substitution rates

The results of the codon preference analysis are shown in Fig. 3a and Table S7. Overall, *P. kingianum* with different putative ploidy levels exhibits the same codon preference. Although the RSCU of AUG and UGG are equal to 1, indicating no usage bias, there are 32 codons with RSCU less than 1 and another 32 codons with RSCU greater than 1 (Fig. 3a; Table S7). Additionally, the RSCU of 30 codons in putative diploid *P. kingianum* are higher than that putative tetraploid *P. kingianum* (Fig. 3a; Table S7), while for 31 codons, the RSCU in putative tetraploid *P. kingianum* are higher than those in putative diploid *P. kingianum* (Fig. 3a; Table S7).

Further comparison of nucleotide substitution rates between putative diploid *P. kingianum* and putative tetraploid *P. kingianum* revealed that although the values of Ka, Ks, and Ka/Ks showed no significant differences ($P > 0.05$; Fig. 3b, c, d), putative tetraploid *P. kingianum* exhibited higher values than the putative diploid *P. kingianum* (Fig. 3b).

Prediction of RNA editing events

Using PREPACT3 to predict RNA editing events within the mitochondrial protein coding genes of *P. kingianum*, it was found that the number of RNA editing sites is very

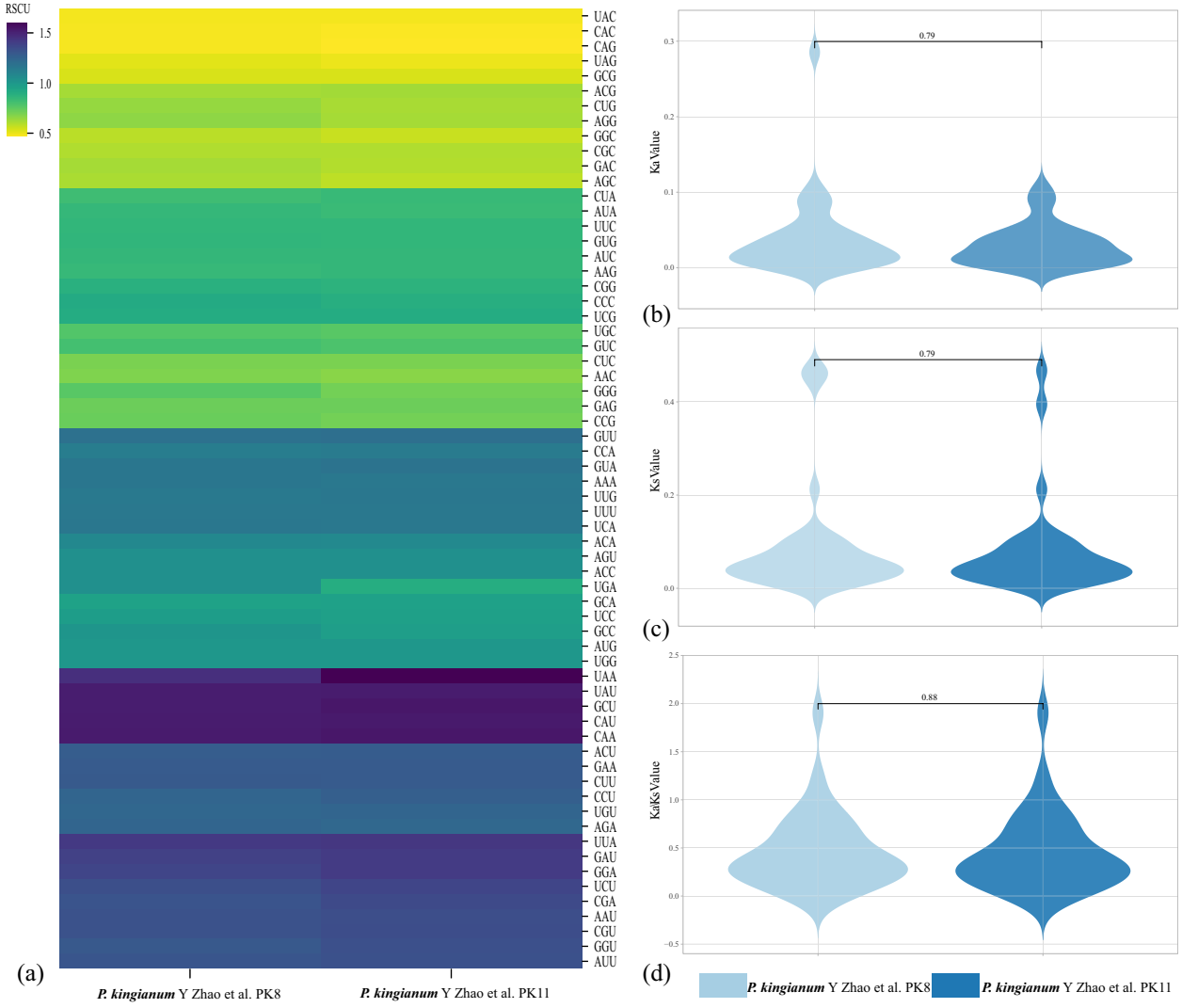

Fig. 3 Codon usage bias and evolutionary pressure of mitochondrial protein coding gene homologues in different putative ploidy levels of *Polygonatum kingianum*. **a** Analysis of relative synonymous codon usage (RSCU). **b** Values of Ka. **c** Values of Ks. **d** Values of Ka/Ks

similar between the putative ploidy levels (Fig. 4; Table S8). Specifically, putative diploid *P. kingianum* has 709 RNA editing sites, while putative tetraploid *P. kingianum* has 710 (Fig. 4; Table S8). Among all RNA editing sites, C-to-U editing sites dominate across the different putative ploidy levels of *P. kingianum*, ranging from 599 in putative tetraploid *P. kingianum* to 602 in putative diploid *P. kingianum*, whereas U-to-C editing sites range only from 107 in putative diploid *P. kingianum* to 111 in putative tetraploid *P. kingianum* (Fig. 4; Table S8). Of the 37 mitochondrial protein coding genes in the different putative ploidy levels of *P. kingianum*, *nad*4 has the highest number of RNA editing sites (56), followed by *ccm*FN (51) and *mtt*B (40) (Fig. 4; Table S8). In both putative ploidy types of *P. kingianum*, *rps*11 has only one RNA editing site (Fig. 4; Table S8). Furthermore, except for *nad*6, *nad*9, *rps*1, and *rps*11, which exhibit only C-to-U editing, all other genes contain both C-to-U and U-to-C RNA editing sites (Fig. 4; Table S8). Notably, a key difference in RNA editing sites between the ploidy types lies in the number of C-to-U edits in specific genes: putative tetraploid *P. kingianum* has 15 and 24 C-to-U editing sites in *cox*2 and *nad*3, respectively, whereas putative diploid *P. kingianum* has 30 and 12 C-to-U editing sites in *cox*2 and *nad*3, respectively (Fig. 4; Table S8).

Mitochondrial-based phylogeny

The alignment of the 41 mitochondrial protein coding genes was 34,141 characters long, of which 5,963 characters were parsimony-informative. The GTR + F + I + G4 was determined as the best-fit model for both ML and BI analyses. The phylogenetic tree resulting from the dataset is shown in Fig. 5. All seven monophyletic orders were recovered with strong support (Fig. 5). The Magnoliales

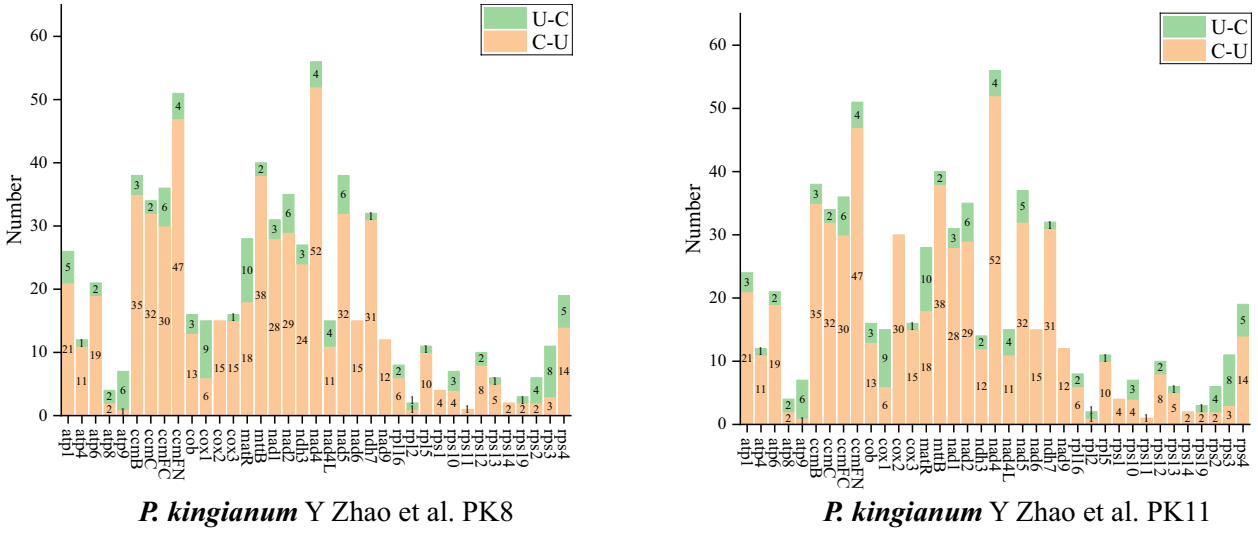

Fig. 4 Prediction of RNA editing sites in the protein coding genes of different putative ploidy levels of *Polygonatum kingianum* mitogenomes

clade was strongly supported as the second branching lineage (ML-BS = 100, BI-PP = 1.00; Fig. 5), followed by Alismatales, Asparagales, Pandanales, Arecales, and Poales. Except for the sister relationship of Arecales and Poales (ML-BS = 55, BI-PP = 0.65), the remaining relationships were received strong support (ML-BS ≥ 80, BI-PP ≥ 0.90; Fig. 5).

Within the Asparagales clade, the two individuals of *P. kingianum* formed a monophyletic clade (ML-BS = 100, BI-PP = 1.00; Fig. 5) and were sister to the remaining species of Asparagales. *P. kingianum* was recovered closer relationship to *Chlorophytum comosum* (ML-BS = 85, BI-PP = 0.95; Fig. 5). It is worth mentioning that the branch lengths of *P. kingianum* individuals with different putative ploidy levels are not significant (Fig. 5).

Discussion

This study assembled the organelle genomes of *P. kingianum* with different putative ploidy levels using both short-read and long-read sequencing (Figs. 1, S1; Table S2). Comparative analyses revealed that the chloroplast genomes are highly stable across putative ploidy levels, showing no significant variation in structure or repetitive sequence characteristics (Figs. 1d, e, S2; Tables S4-S6). In contrast, although from the same species, the mitogenomes differ markedly between putative ploidy types. Unlike the chloroplast genomes displayed the typical quadripartite structure (Fig. 1d, e), the mitogenomes of different putative *P. kingianum* ploidies exhibit a multi-branched linear structure (Figs. 1a, b, c, f, S1), which is associated with a high abundance of repetitive sequences (Fig. 2; Tables S4-S6). The mitogenome of putative tetraploid *P. kingianum* consists of two chromosomes, whereas that of putative diploid *P. kingianum* comprises only one chromosome (Figs. 1a, b, c, f, S1; Table 1). While this finding is largely consistent with organelle genome patterns observed in other plants, it presents certain differences that enhance our understanding of plant mitochondrial genome evolution. For instance, the assembly of the mitogenomes of *Tinospora sagittata* (ca. 513 kb) revealed repeat-driven recombination events and codon usage bias, providing a valuable model for studying the evolution of medicinal plants [17]. However, the mitogenome of *T. sagittata* consists of a single chromosome, but structural variations including one translocation and two inverted segments exist between different putative ploidies. In the model plant *Arabidopsis thaliana*, although chloroplast genomes from different individuals similarly possess a single quadripartite structure, pan-mitogenomic studies have identified at least three distinct mitochondrial genome structures, primarily shaped by repetitive sequences [25]. Tao et al. [26] also had been detected that mitogenome recombination of *P. kingianum* were mediated by repeats. Therefore, we also propose that in plant, even within the same species, the mitogenome may exhibit diverse structural conformations mediated by the amount of repeat. Besides, with decoding the mitogenomes of *P. sibiricum* [22] and *P. kingianum* not only fill the molecular genetic knowledge gap, but also offers comparative genomic insights into their evolutionary relationships with close relatives.

The *nad* gene family plays a crucial role in the respiratory process [48]. Compared to the putative diploid *P. kingianum*, *nad*3 and *nad*5 genes have undergone duplication in putative tetraploid *P. kingianum* (Fig. 1, a, b, c, f; Table S3). Copy number variation of *nad*3 is observed across species, for instance, maize (*Zea mays* [AY506529]) [49] has a single copy, while carrot (*Daucus carota* subsp. *sativus* [NC017855]) [50] contains three copies. The variation in gene copy number among

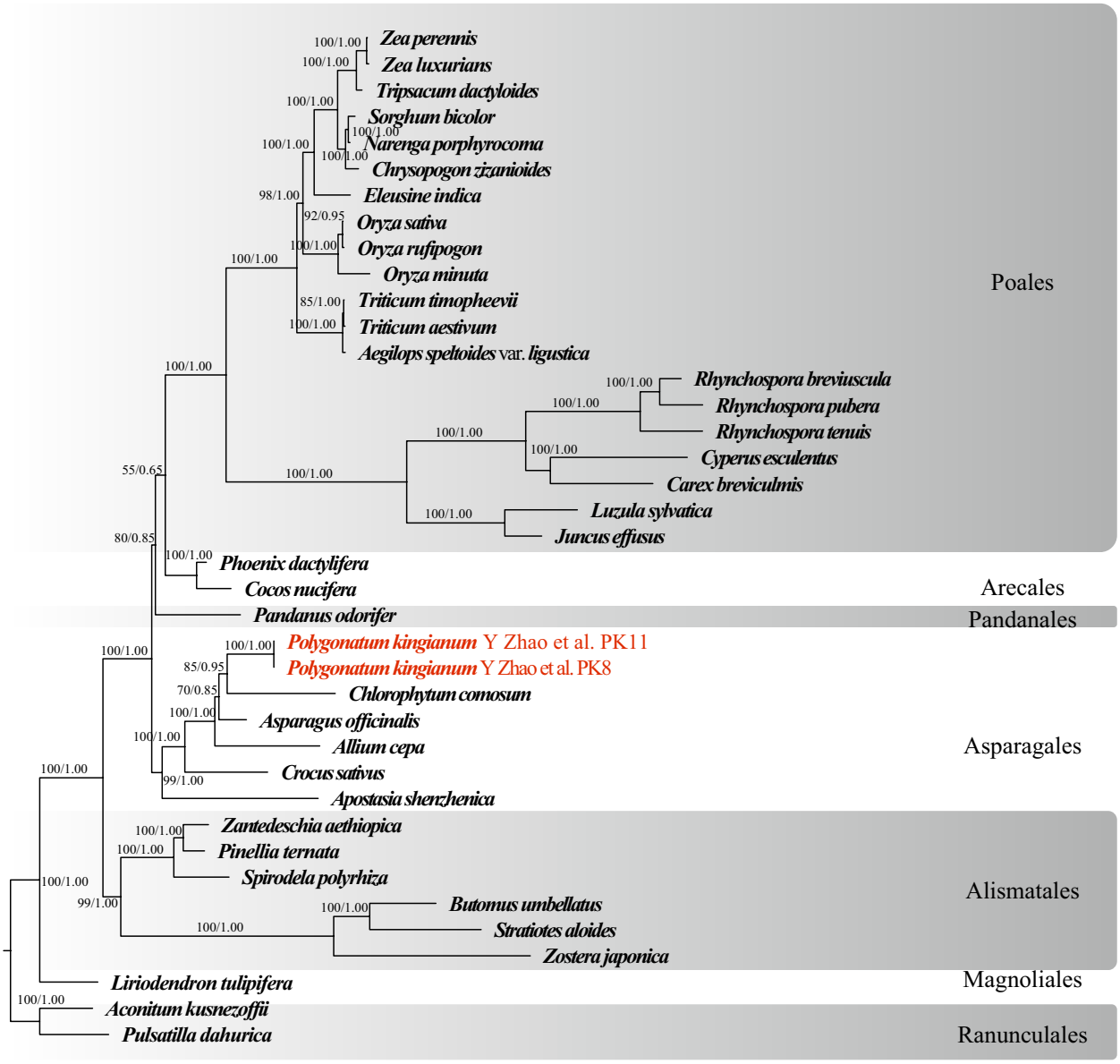

Fig. 5 Maximum likelihood tree based on 41 mitochondrial protein coding genes. Maximum Likelihood Bootstrap Support values (ML-BS) and Bayesian Inference Posterior Probability (BI-PP) are given above the branches

different putative *P. kingianum* ploidy levels may not only relate to respiration but could also confer better adaptability to various stress conditions due to the presence of these additional copies.

Whole-genome duplication is considered a primary driver of plant polyploidization [51, 52]. Following polyploidization, gene duplication occurs between species of different ploidy levels, accompanied by changes in repetitive sequences [53]. Although the mitogenome sizes in *P. kingianum* do not scale proportionally with putative ploidy, a correlation exists between repetitive sequences and ploidy (Figs. 2, S2; Tables S4-S6). Specifically, the number of SSRs in putative tetraploid *P. kingianum* (226 SSRs in the mitogenome, 202 SSRs in the chloroplast genome) is nearly double that in putative diploid *P. kingianum* (105 SSRs in the mitogenome, 159 SSRs in the chloroplast genome; Fig. 2, Table S4), while the number of TRs in the putative diploid mitogenome (26 TRs) is 1.5 times that of the putative tetraploid (17 TRs). No significant differences were observed for TRs and DRs in the chloroplast genomes between ploidies (Fig. 2; Table S4-S6). This difference is not only related to the more complex structural variation of the mitogenomes compared to the chloroplast genomes, we speculate that it may represent a mechanism by which mitogenomes maintain uniparental inheritance through a high abundance of repetitive sequences during polyploidization, but need further studies. In most plants,

both chloroplast and mitochondrial genomes are maternally inherited and thus less susceptible to factors such as hybridization [54]. Additionally, nuclear-encoded organelle-targeted genes (such as *msh*1, *OSB*, *Why*, and *Rec*A) play significant roles in the replication, recombination, and repair of organellar genomes [25, 55–57]. These nucleus-encoded targeted proteins may exhibit varying expression patterns, which could lead to differences in the relationship between ploidy and repetitive sequences in organellar genomes.

RNA editing is a common phenomenon in organelles, particularly within mitogenomes, and is crucial for protein function and gene expression [58–60]. Although the total number of RNA editing sites is similar across different *P. kingianum* ploidy levels, distinct editing patterns exist for *cox*2 and *nad*3 (Fig. 4; Table S8). Specifically, the putative tetraploid *P. kingianum* has twofold the number of editing sites in *nad*3 compared to the putative diploid. The increase of RNA editing sites in *nad*3 between ploidy levels may be associated with differential drought adaptation [61], salt tolerance [62], and other stress responses. We advocate integrating more other evidence in future which will be helpful to understand the adaptive mechanisms that may exist behind this. In addition, differences in the number of RNA editing sites in *cox*2 might be associated with varying cold stress adaptation [63], but further research is needed. Furthermore, phylogenetic trees constructed using mitochondrial genes often show high support values, but our mitochondrial-based phylogeny did not reveal significant speciation between the putative ploidy levels of *P. kingianum* (Fig. 5). The minimal differences in nucleotide substitution rates further support the conclusion that the morphologically variable *P. kingianum* likely constitutes only one species [1]. Although Wang et al. [16] found that *P. kingianum* with different flower colors can be divided into at least three clades based on the chloroplast genomes, the branch lengths in the phylogenetic tree are very short, indicating no significant differentiation among these clades. We propose that morphological variations, such as flower color, may be shaped by environmental heterogeneity [16]. Given that the evolutionary process of *Polygonatum* is predominantly characterized by incomplete lineage sorting, with limited hybridization and introgression [4], we advocate for future research to expand sampling efforts and integrate more evidences for more robust species delimitation. In summary, our work establishes essential genomic resources for molecular authentication and breeding programs while providing novel insights into mitogenome evolution in medicinal plants through comparative genomic analyses.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-026-08828-9.

Supplementary Material 1. Figure S1. Maps of mitochondrial genomes assembly of different putative ploidy levels of *Polygonatum kingianum*. (a) The map of the preliminary assembly of the mitogenome of putative diploid *P. kingianum*. (b) The structure of the mitogenome of putative diploid *P. kingianum* after debranching the initial assembly. (c) The map of the preliminary assembly of the mitogenome putative diploid *P. kingianum*. (d) The structure of the mitogenome of putative diploid *P. kingianum* after debranching the initial assembly.

Supplementary Material 2. Figure S2. Repeat sequences analyses of different putative ploidy levels of *Polygonatum kingianum* chloroplast genomes. (a) The frequency of identified simple sequence repeats. (b) Statistics of tandem repeats sequence with different lengths. (c) Type and quantity of simple sequence repeat motifs. (d) Statics of dispersed repeat sequence types.

Supplementary Material 3. Table S1. The list of the mitogenomes used to construct phylogenetic trees. Table S2. Summary of the flow cytometer and sequence results. Table S3. Gene composition of the mitogenome. Table S4. The frequency of identified SSRs. Table S5. Statics of tandem repeats. Table S6. Statics of dispersed repeat sequences. Table S7. Codon usage bias analysis of different putative ploidy levels of *Polygonatum kingianum* mitogenomes. Table S8. RNA editing sites identified in the mitogenomes of different putative ploidy levels of *Polygonatum kingianum*.

Acknowledgements

The authors also thank three anonymous reviewers and editors for their careful review to improve the manuscript.

Authors' contributions

Y.Z., Y.X.L., and Y.Y.Z. contributed to the study conception and design. Material preparation was performed by M.X.Z. and L.H.W. Experiments and data analyses were performed by J.Z., Z.H.C., and J.G.W. Data collection and interpretation were performed by J.Z., Z.H.C. and X.Y.Y. The first draft of the manuscript was written by J.Z. and Z.H.C.

Funding

This study was supported by grants from Yunnan Fundamental Research Projects (Grant No. 202501AT070075), the Yunnan Fundamental Research Projects (Grant No. 202401BD070001-055), the Yunnan Provincial Major Science and Technology Special Program (Grant No. 202502AS100013), and the Agricultural Joint Program Key Project of Yunnan Province (Grant No. 202301BD070001-205).

Data availability

The new sequenced organelle genome sequences supporting the conclusions of this article are available in GenBank (https://www.ncbi.nlm.nih.gov/) with accession numbers: PX777949, PX777950, PX777962, PX777963, and PX777964, respectively. The Illumina data had been deposited in the NCBI Sequence Read Archive (SRR36644920, SRR36644921). Nanopore data and materials provided in this manuscript are available from the corresponding author upon reasonable request. The new sequenced organelle genome sequences [including both "fasta" and "gb" files], multiple sequence alignments, concatenated alignments, and phylogenetic trees for this study are publicly available at Figshare repository: https://doi.org/10.6084/m9.figshare.30945248.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details
[1]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, College of Plant Protection, Yunnan Agricultural University, Kunming 650201, China
[2]Medicinal Plants Research Institute, Yunnan Academy of Agricultural Sciences, Kunming 650200, China
[3]School of Ecology and Environmental Science, Yunnan University, Kunming 650504, Yunnan, China
[4]School of Life Sciences, Yunnan University, East Outer Ring Road, Chenggong District, Kunming 650500, Yunnan, China

Received: 24 December 2025 / Accepted: 20 April 2026
Published online: 29 April 2026

References

1. Chen XQ, Tamura MN. Polygonatum. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China, vol. 7. Beijing and St. Louis: Science Press and Missouri Botanical Garden Press; 2000. pp. 223–32.
2. Meng Y, Nie ZL, Deng T, Wen J, Yang YP. Phylogenetics and evolution of phyllotaxy in the Solomon's seal genus *Polygonatum* (Asparagaceae: Polygonateae). Bot J Linn Soc. 2014;176:435–51.
3. Wang J, Qian J, Jiang Y, Chen X, Zheng B, Chen S, et al. Comparative analysis of chloroplast genome and new insights into phylogenetic relationships of *Polygonatum* and tribe Polygonateae. Front Plant Sci. 2022;13:882189.
4. Qin YQ, Zhang MH, Yang CY, Nie ZL, Wen J, Meng Y. Phylogenomics and divergence pattern of *Polygonatum* (Asparagaceae: Polygonateae) in the north temperate region. Mol Phylogenet Evol. 2024;190:107962.
5. Zhao P, Zhao C, Li X, Gao Q, Huang L, Xiao P, et al. The genus *Polygonatum*: a review of ethnopharmacology, phytochemistry and pharmacology. J Ethnopharmacol. 2018;214:274–91.
6. Committee NP. Pharmacopoeia of the people's Republic of China (Volume I). Beijing: China Medical Science and Technology; 2020.
7. Zhao X, Li J. Chemical constituents of the genus *Polygonatum* and their role in medicinal treatment. Nat Prod Commun. 2015;10:683–8.
8. Jin J, Lao J, Zhong C, He W, Xie J, Hu G, et al. Complete chloroplast genome of a medicinal species *Polygonatum kingianum* in China (Asparagaceae, Asparagales). Mitochondrial DNA B Resour. 2020;5(1):959–60.
9. Wicke S, Schneeweiss GM, de Pamphilis CW, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 2011;76:273–97.
10. Xue TT, Janssens SB, Liu BB, Yu SX. Phylogenomic conflict analyses of the plastid and mitochondrial genomes via deep genome skimming highlight their independent evolutionary histories: a case study in the cinquefoil genus *Potentilla sensu lato* (Potentilleae, Rosaceae). Mol Phylogenet Evol. 2024;190:107956.
11. Zhao J, Liang ZL, Fang SL, Li RJ, Huang CJ, Zhang LB, et al. Phylogenomics of *Paragymnopteris* (Cheilanthoideae, Pteridaceae): insights from plastome, mitochondrial, and nuclear datasets. Mol Phylogenet Evol. 2024;204:108253.
12. Feng Y, Xiang X, Akhter D, Pan R, Fu Z, Jin X. Mitochondrial phylogenomics of Fagales provides insights into plant mitogenome mosaic evolution. Fronti Plant Sci. 2021;12:762195.
13. Lin D, Shao B, Gao Z, Li J, Li Z, Li T, et al. Phylogenomics of angiosperms based on mitochondrial genes: insights into deep node relationships. BMC Biol. 2025;23(1):45.
14. Nie LY, Wang J, Huang L, Kong JL, Nie B, Tembrock LR, Dong SS, Tiwari R, Wang H, Kan SL, Zou XH, Wu ZQ. Dissecting mitogenomic conflict to illuminate angiosperm deep phylogeny: sequence and architectural evidence. Plant Divers. 2025;48(1):41–58. https://doi.org/10.1016/j.pld.2025.10.003.
15. Tang YC. *Polygonatum* Mill. In: Wang FT, Tang T, editors. Flora Reipublicae Popularis Sinicae. Beijing: Science Press; 1978. p. 65.
16. Wang J, Guo X, Li B, Zhang C, Yi Y, Tang M, et al. Molecular identification and phylogenetic analysis of *Polygonatum kingianum* with different floral colors on the basis of chloroplast genomes. BMC Plant Biol. 2025;25(1):1285.
17. Zhao J, Chen ZH, Huang PC, Chen LW, Zhang MX, Wang LH, et al. Complete mitogenome characteristics and phylogenetic analysis of traditional Chinese medicinal plant *Tinospora sagittata* (Oliv.) Gagnep. from the Menispermaceae family. BMC Plant Biol. 2025;25(1):165.
18. Tang JY, Wei R, Zhang XC, Xiang QP. Mitogenome-based phylogenomics provides insights into the positions of the enigmatic sinensis group and the sanguinolenta group in Selaginellaceae (Lycophyte). Mol Phylogenet Evol. 2023;179:107673.
19. Liang H, Qi H, Chen J, Wang Y, Liu M, Sun X, et al. Assembly and analysis of the first complete mitochondrial genome sequencing of main Tea-oil *Camellia* cultivars *Camellia drupifera* (Theaceae): revealed a multi-branch mitochondrial conformation for *Camellia*. BMC Plant Biol. 2025;25(1):13.
20. Gao K, Guo T, An X. Comprehensive analysis of the multi-rings mitochondrial genome of *Populus tomentosa*. BMC Genomics. 2025;26(1):23.
21. Zhao J, Yu H, Rothfels CJ, Yang JJ, Fang SL, He ZR, et al. Widespread incongruence in the phylogenomics of the ancient land plant lineage, Selaginellaceae (lycophytes). Mol Phylogenet Evol. 2026;214:108454.
22. Liu M, Fan R, Wang C, Dai L, Chu S. Complete analysis and phylogenetic analysis of *Polygonatum sibiricum* mitochondria. BMC Plant Biol. 2025;25(1):471.
23. Bendich AJ. Reaching for the ring: the study of mitochondrial genome structure. Curr Genet. 1993;24(4):279–90.
24. Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, Michelmore RW, et al. The alternative reality of plant mitochondrial DNA: one ring does not rule them all. PLoS Genet. 2019;15(8):e1008373.
25. Zou Y, Zhu W, Hou Y, Sloan DB, Wu ZQ. The evolutionary dynamics of organellar pan-genomes in *Arabidopsis thaliana*. Genome Biol. 2025;26:240.
26. Tao A, Tao J, Gan Z, Zhang Y, Yin C, Jiang G. The characterisation of the complete mitochondrial genome of *Polygonatum kingianum* reveals recombination mediated by repeats associated with DNA replication. Sci Rep. 2025;15(1):19980.
27. Dolezel J, Greilhuber J, Suda J. Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc. 2007;2(9):2233–44.
28. Doyle JJ. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.
29. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90.
30. Jin JJ, Yu WB, Yang JB, Song Y, de Pamphilis CW, Yi TS, et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020;21(1):241.
31. Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, et al. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 2019;47:W65–73.
32. He W, Xiang K, Chen C, Wang J, Wu Z. Master graph: an essential integrated assembly model for the plant mitogenome based on a graph-based framework. Brief Bioinform. 2023;24(1):bbac522.
33. Bi C, Shen F, Han F, Qu Y, Hou J, Xu K, et al. PMAT: an efficient plant mitogenome assembly toolkit using low-coverage HiFi sequencing data. Hortic Res. 2024;11(3):uhae023.
34. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 2015;31:3350–2.
35. Chan PP, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol. 2019;1962:1–14.
36. Lohse M, Drechsel O, Kahlau S, Bock R. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. 2013;41:W575–81.
37. Beier S, Thiel T, Münch T, Scholz U, Mascher M. Misa-web: a web server for microsatellite prediction. Bioinformatics. 2017;33(16):2583–5.
38. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. Reputer: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29(22):4633–42.
39. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80.
40. Zhang D, Gao F, Jakovlic I, Zou H, Zhang J, Li WX, Wang GT. PhyloSuite an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 2020;20:348–55.
41. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteomics Bioinformatics. 2010;8(1):77–80.
42. Lenz H, Hein A, Knoop V. Plant organelle RNA editing and its specificity factors: enhancements of analyses and new database features in PREPACT3.0. BMC Bioinformatics. 2018;19(1):255.
43. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.
44. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972–3.

45. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
46. Hoang DT, Chernomor O, Haeseler AV, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35(2):518–22.
47. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–42.
48. Hashida SN, Takahashi H, Uchimiya H. The role of NAD biosynthesis in plant development and stress responses. Ann Bot. 2009;103(6):819–24.
49. Clifton SW, Minx P, Fauron CM, Gibson M, Allen JO, Sun H, et al. Sequence and comparative analysis of the maize NB mitochondrial genome. Plant Physiol. 2004;136(3):3486–503.
50. Iorizzo M, Senalik D, Szklarczyk M, Grzebelus D, Spooner D, Simon P. De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. BMC Plant Biol. 2012;12:61.
51. Clark JW, Donoghue PCJ. Whole-genome duplication and plant macroevolution. Trends Plant Sci. 2018;23(10):933–45.
52. Soltis PS, Soltis DE. Ancient WGD events as drivers of key innovations in angiosperms. Curr Opin Plant Biol. 2016;30:159–65.
53. Schranz ME, Mohammadin S, Edger PP. Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model. Curr Opin Plant Biol. 2012;15(2):147–53.
54. Wang J, Kan S, Liao X, Zhou J, Tembrock LR, Daniell H, et al. Plant organellar genomes: much done, much more to do. Trends Plant Sci. 2024;29(7):754–69.
55. Sloan DB, Taylor DR. Evolutionary rate variation in organelle genomes: the role of mutational processes. In: Bullerwell C, editor. Organelle Genetics. Berlin: Springer; 2012. p. 123–46.
56. Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 2004;5(2):123–35.
57. Kleine T, Maier UG, Leister D. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Annu Rev Plant Biol. 2009;60:115–38.
58. Rurek M, Szklarczyk M, Adamczyk N, Michalik B, Augustyniak H. Differences in editing of mitochondrial *nad*3 transcripts from CMS and fertile carrots. Acta Biochim Pol. 2001;48(3):711–7.
59. Takenaka M, Zehrmann A, Verbitskiy D, Härtel B, Brennicke A. RNA editing in plants and its evolution. Annu Rev Genet. 2013;47:335–52.
60. He P, Xiao G, Liu H, Zhang L, Zhao L, Tang M, et al. Two pivotal RNA editing sites in the mitochondrial *atp*1 mRNA are required for ATP synthase to produce sufficient ATP for cotton fiber cell elongation. New Phytol. 2018;218:167–82.
61. Yuan H, Liu D. Functional disruption of the pentatricopeptide protein SLG1affects mitochondrial RNA editing, plant development, and responses to abiotic stresses in *Arabidopsis*. Plant J. 2012;70:432–44.
62. Ramadan AM. Salinity effects on *nad*3 gene RNA editing of wild barley mitochondria. Mol Biol Rep. 2020;47:3857–65.
63. Kurihara-Yonemoto S, Handa H. Low temperature affects the processing pattern and RNA editing status of the mitochondrial *cox*2 transcripts in wheat. Curr Genet. 2001;40(3):203–8.

Publisher's Note