BMC Genomics

RESEARCH Open Access

Chloroplast genome comparison and taxonomic reassessment of *Polygonatum sensu Lato* (Asparagaceae): implications for molecular marker development in traditional medicinal plants

Yingfeng Hu[1], Siqin Wang[1], Zhenzhen Xu[1], Siyu Yan[1], Maroof Ali[1,4], Zhizhong Li[2*] and Jianwen Shao[1,3*]

Abstract

The Polygonati Rhizoma have generated significant market attention for their medicinal and culinary applications. However, morphological similarities and ambiguous species boundaries complicate the identification of genera and species, thereby impeding product development and utilization within *Polygonatum sensu lato*. Despite the widespread application of the chloroplast genome for taxonomic boundary revisions for *Polygonatum s.l.*, a critical gap persist regarding their genomic applicability and the lack of standardized pipelines for developing species-specific molecular markers capable of rapid discrimination among species. This study aims to assess the effectiveness of chloroplast genomes in clarifying the current taxonomic status of the genera and species of *Polygonatum s.l.*, and develop a reliable process for rapid identification of designated species from other species. A total of 21 chloroplast genomes were sequenced and assembled, and subsequent analyses included phylogenetic inference, multiple molecular species delimitation methods, and an automated screening framework were employed for subsequent analysis. Comparative analyses revealed relatively conserved chloroplast genomes, with notable variation limited primarily to the length of IR and LSC regions. By integrating multiple delimitation methods, the chloroplast genome validated 82.46% of the current classifications of *Polygonatum s.l.*, demonstrating strong support (90.63%) for species represented by multiple sequences, yet only moderate support (70%) for those with single-sequence representation. Additionally, this study established and validated a scalable molecular marker development framework, spanning from identification of species-specific SNPs/InDels to the design of high-resolution molecular markers, illustrated through case studies involving *Heteropolygonatum* and three medicinally significant *Polygonatum* species.

Keywords DNA barcoding, Polygonati rhizoma, Molecular marker, *Heteropolygonatum*

*Correspondence:
Zhizhong Li
lizhizhong@ahnu.edu.cn
Jianwen Shao
shaojw@ahnu.edu.cn

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

Introduction

Polygonatum sensu lato (*s.l.*), comprising *Polygonatum* Mill. and *Heteropolygonatum* M. N. Tamura & Ogisu, includes over 80 species distributed predominantly across temperate regions of Eurasia and North America [1–5]. Several species, notably *P. sibiricum*, *P. cyrtonema*, and *P. kingianum*, possess significant economic value as multipurpose resources in traditional medicine [6, 7]. The rhizomes (Polygonati Rhizoma) of these species are rich in bioactive compounds, including polysaccharides, saponins, and flavonoids [8, 9]. Furthermore, these rhizomes also exhibit a wide spectrum of pharmacological activities, such as modulation of intestinal microbiota, immune enhancement, antioxidant, anti-inflammatory, antidiabetic, antitumor, hypolipidemic, anti-fatigue, neuroprotective, and osteoporosis-preventing effects [8, 10, 11]. These therapeutic attributes, coupled with their historical utilization in Chinese herbal formulations for over two millennia, have driven modern applications ranging from nutraceutical development to innovative processing techniques like the traditional 'nine-steam-nine-bask' method [7, 12].

Accurate species identification is essential for ensuring the quality of economic crops, conserving germplasm resources, supporting industrial applications, and advancing scientific research [13–15]. Currently, species delimitation within *Polygonatum s.l.* predominantly relies on morphological characterization of diagnostic traits such as rhizome, phyllotaxis, and inflorescence [1]. However, this method of morphological identification remains challenging because many scenarios rely solely on *Polygonatum* rhizomes, complicating quality control in herbal product manufacturing [16, 17]. These persistent challenges have led to adjustments in species boundaries for more than seven species between *Polygonatum* and *Heteropolygonatum* in the past two decades, resulting in substantial taxonomic ambiguities [18–22]. Diagnostic features differentiating these two genera are challenging to observe directly. *Polygonatum* is characterized by a chromosome number of x = 9–15 and valvate tepals, while *Heteropolygonatum* exhibits chromosome numbers of x = 16, 32 with imbricate tepals [18, 20]. Previous phylogenetic studies employing multi-locus chloroplast markers have effectively resolved generic-level delimitation between *Polygonatum* and *Heteropolygonatum*, and have identified three major lineages within *Polygonatum* [23, 24]. However, these methods have demonstrated limited resolution at the species level. Recent advances in phylogenomics and phylogeographic analyses utilizing whole-chloroplast genomes have demonstrated greater potential in resolving species boundaries [2, 25, 26]. Crucially, whole-chloroplast genomes enable comprehensive detection of structural variants and provide an increased number of variable sites, thereby offering unprecedented resolution at the species level [17]. However, critical gaps persist in translating these evolutionary insights and their application potential into practical frameworks for authenticating medicinal species [27, 28]. This gap underscores the urgent need to bridge phylogenetic frameworks and practical identification tools, particularly for high-value species within *Polygonatum*, including the three traditional medicinal species [17, 29, 30].

The past decade has witnessed a surge in genetic methods for species identification, laying a critical foundation for systematically evaluating the utility of chloroplast genomes for clarifying current taxonomic classifications [31]. Current validation frameworks predominantly rely on single-locus datasets and comparative analyses of genetic distances, supplemented by multiple molecular species delimitation methods [32–34]. However, divergent taxonomic conclusions frequently occur due to data heterogeneity (e.g., cytonuclear discordance) and variable sampling densities across studies [28, 35]. These discrepancies highlight the imperative need for evaluating the taxonomic applicability of chloroplast genomes. Furthermore, empirical comparisons of these delimitation algorithms underscore the necessity of integrative validation frameworks to quantitatively assess support from molecular datasets for current species boundaries [31].

The integration of rigorously validated genomic datasets facilitates the development of species-specific molecular markers, significantly enhancing efficiency in species identification for morphologically cryptic lineages. These markers, including InDel regions and species-specific sites resulting from single nucleotide polymorphisms (SNPs), insertions or deletions at identical genomic loci across related species [36], are effective for species identification and combating herbal adulteration in commercial markets [37, 38]. For example, the development of InDel primers for *Cynanchum wilfordii* and *C. auriculatum* has effectively reduced adulteration of herbal materials [39], while species-specific primers for *Arnebiae* Forssk. species have enhanced identification accuracy [40]. Combining chloroplast genome analysis with these two types of species-specific markers offers a more versatile and reliable species identification approach, providing a foundation for quality control and sustainable cultivation of economically important species. This integrated strategy ensures taxonomic accuracy and enhances both scientific research and practical industrial applications. In this study, we sequenced 21 complete chloroplast genomes from *Polygonatum s.l.*, prioritized to augment underrepresented groups, and integrated 102 publicly available datasets (comprising 11 from *Heteropolygonatum*, 90 from *Polygonatum*, and one outgroup) from public databases. Collectively, these sequences represent all three major clades within *Polygonatum* (sect. *Verticillata*, sect. *Sibirica*, and sect. *Polygonatum*), species phylogenetically

closely related to the three medicinal *Polygonatum* species, and additional intraspecific variation. The objectives of this study were: (1) to identify patterns of chloroplast genome variation in *Polygonatum s.l.*; (2) to evaluate the effectiveness of chloroplast genome in delineating the currently accepted species of *Polygonatum s.l.* using various molecular species delimitation programs; and (3) to explore the distribution of species-specific sites for developing and validating specific primers to facilitate accurate genus and traditional medicinal species identification. Our study provides a comprehensive evaluation of chloroplast genomes, offering new insights and methodological frameworks for resolving taxonomic uncertainties at various levels in key medicinal plant species.

Materials and methods

Plant material and DNA sequencing

From 2020 to 2023, fieldwork was conducted in Anhui, Guizhou, Shandong, Shaanxi, Sichuan, Yunnan, and Zhejiang Provinces, China, where samples were collected from 21 populations of 13 species belonging to *Polygonatum s.l.* (Table 1; Fig. 1). These samples were identified by Yingfeng Hu and Prof. Jianwen Shao based on morphological characterization, and were deposited at the herbarium of Anhui Normal University (ANUB). One population of *Heteropolygonatum*, morphologically similar to *H. alternicirrhosum* (Hand.-Mazz.) Floden, but distinguished by coarse hairs on its stems and peduncles, was temporarily designated as *Heteropolygonatum* sp. Kangding in this study (Table 1). To evaluate chloroplast genome efficacy in species delimitation via genetic distance analysis, targeted sampling was conducted (Table 1) for taxa in *Polygonatum s.l.* with insufficient genomic data, ensuring at least two sequences per species, and prioritizing underrepresented groups. This optimized dataset enhanced phylogenetic resolution while expanding the chloroplast genomic repository, thus strengthening the validity of species boundary determination [17].

Fresh leaves from several individuals in each population were collected and dried in silica gel for subsequent molecular analysis. Total genomic DNA was extracted using the Tiangen DNA-secure Plant Kit (DP320) according to the manufacturer's protocol, and DNA quality was evaluated using 2% agarose gel electrophoresis. For each sample, 350-bp DNA libraries were prepared using the MagicSeq DNA Library Prep Kit (M319), followed by sequencing with 150-bp paired-end reads on the BGI DNBSEQ-T7 platform at the Germplasm Bank of Wild Species in Southwest China (GBOWS, Kunming, China). After filtering low-quality reads with Fastp v0.23.2 [41]

Table 1 Sampling information of the newly sequenced samples in this study

Scientific name	GenBank number	Location	Longitude	Latitude	Herbarium accession numbers
H. alternicirrhosum	PQ591744	Dagang Village, Kangding City, Sichuan Province	102.1457E	30.0705 N	ANUB100117
H. binatifolium	PQ436974	Baifotai, Fanjing Mountain, Tongren City, Guizhou Province	108.6646E	27.9089 N	ANUB100118
H. binatifolium	PQ436975	Jindaoxia, Fanjing Mountain, Tongren City, Guizhou Province	108.6742E	27.9163 N	ANUB100119
H. binatifolium	PQ591745	Diecuixi, Longcanggou Town, Yaan City, Sichuan Province	102.8919E	29.6163 N	ANUB100120
H. binatifolium	PQ591746	Renshenggou, Yaan City, Sichuan Province	102.8945E	29.6154 N	ANUB100121
Heteropolygonatum. sp. Kangding	PQ436973	Yulin Village, Kangding City, Sichuan Province	101.9613E	29.9677 N	ANUB100122
P. curvistylum	PV055066	Qingfengxia, Baoji City, Shanxi Province	107.4463E	34.0353 N	ANUB100123
P. caulialatum	PV055067	Baiyangwan, Kangding City, Sichuan Province	101.9587E	29.9665 N	ANUB100124
P. cyrtonema var. *gutianshanicuym*	PV055063	Suzhuang Town, Quzhou City, Zhejiang Province	118.1248E	29.1741 N	ANUB100125
P. cyrtonema var. *gutianshanicuym*	PV055053	Hongtan Town, Huangshan City, Anhui Province	117.8611E	30.0911 N	ANUB100126
P. cyrtonema var. *gutianshanicuym*	PV055064	Yanglin Village, Huangshan City, Anhui Province	117.5097E	29.9017 N	ANUB100127
P. franchetii	PV055065	Wulixia, Baoji City, Shanxi Province	107.4367E	33.9881 N	ANUB100128
P. jinzhaiense	PV055061	Yaoluoping, Anqing City, Anhui Province	116.0859E	30.9834 N	ANUB100129
P. jinzhaiense	PV055062	Tiantangzhai, Liuan City, Anhui Province	115.7711E	31.1237 N	ANUB100130
P. langyanese	PV055060	Langya Mountain, Chuzhou City, Anhui Province	118.2866E	32.2778 N	ANUB100131
P. macropodum	PV055059	Mount Tai, Taian City, Shandong Province	117.1428E	36.2995 N	ANUB100132
P. nodosum	PV055057	Yaochiba, Jinfo Mountain, Chongqing City	107.1273E	29.0492 N	ANUB100133
P. nodosum	PV055058	Dujuan Park, Jinfo Mountain, Chongqing City	107.1839E	29.0168 N	ANUB100134
P. punctatum	PV055055	Dawei Mountain, Honghe City, Yunnan Province	103.8806E	22.7381 N	ANUB100135
P. punctatum	PV055056	Dawei Mountain, Honghe City, Yunnan Province	103.8806E	22.7381 N	ANUB100136
P. zanlanscianense	PV055054	Wumei Village, Chizhou City, Anhui Province	117.9712E	30.5829 N	ANUB100137

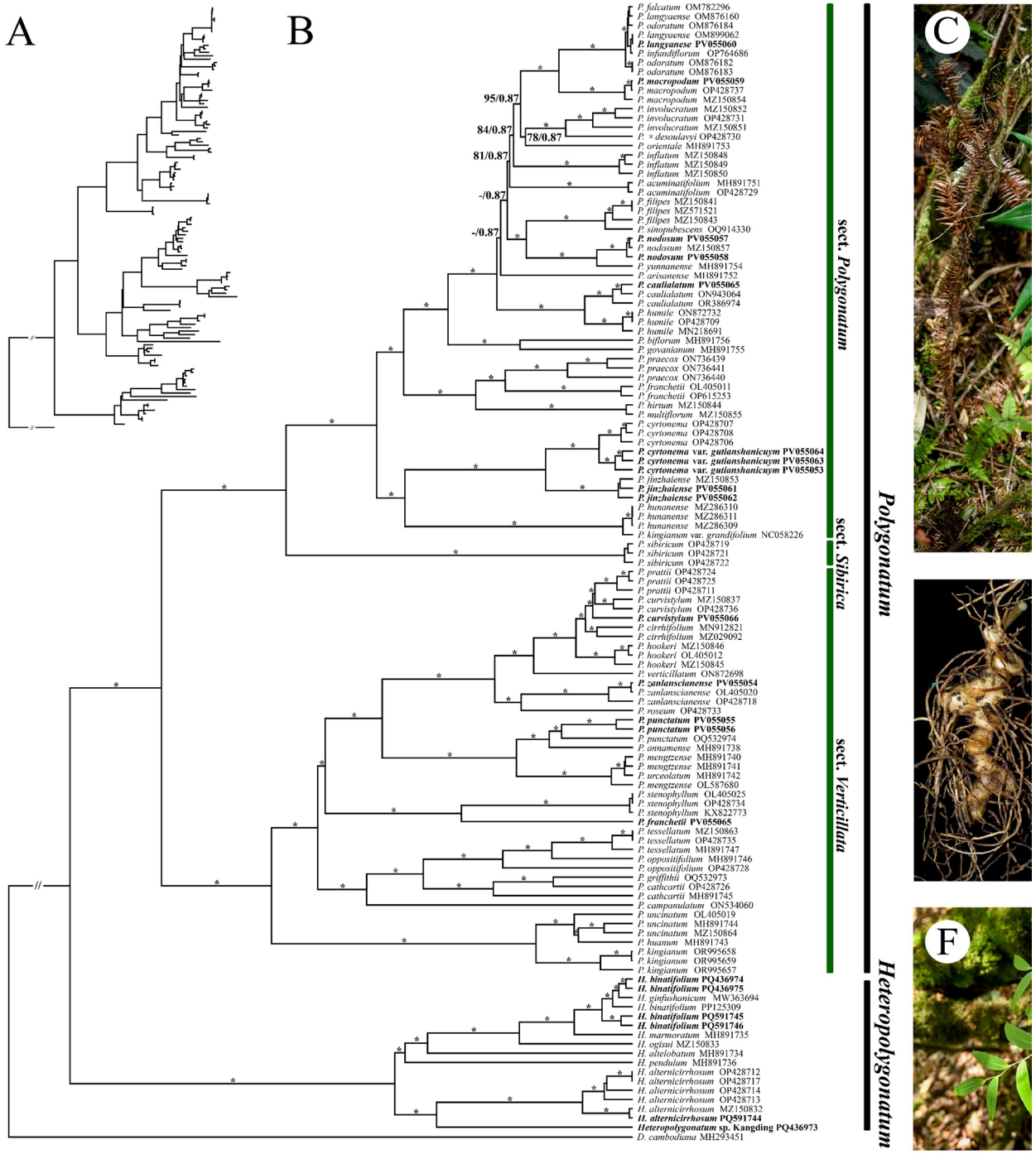

Fig. 1 Phylogenetic relationships trees of *Polygonatum* and *Heteropolygonatum* based on 78 protein-coding genes. Phylogenetic tree with branch lengths based on Maximum Likelihood Method (A). A phylogenetic tree without branch lengths, generated using both Maximum Likelihood and Bayesian Inference (B); Asterisks represent BS = 100 / PP = 1; Branches with no support value; Samples highlighted in bold were newly generated in this study. Habitat (C) and rhizome (D) of *P. cyrtonema*. Rhizome (E) and habitat (F) of *H. binatifolium*

under default parameters, approximately 3 Gb of clean data was generated for each sample.

Complete Chloroplast genome assembly, annotation and comparison analysis

All 21 complete chloroplast genomes were *de novo* assembled using GetOrganelle v1.7.6.1 with *K-mer* values of 21, 85, and 127 [42] while other parameters were kept at default settings. These newly obtained chloroplast genomes, along with those downloaded from public databases, were annotated and re-annotated by PGA [43], with *H. alternicirrhosum* (MZ150832) as a reference. Start and stop codons for each protein-coding gene (PCG) were manually verified and corrected using Geneious v10.2.3 (Biomatters Ltd, Auckland, New Zealand). All newly generated chloroplast genomes were visualized using OGDRAW [44] and deposited in GenBank (https://www.ncbi.nlm.nih.gov/; Table 1).

Integrated with public databases, the chloroplast genomes of 57 species, including seven *Heteropolygonatum* and 50 *Polygonatum*, were analyzed. To identify potential structural variations and sequence lengths discrepancies within the chloroplast genomes, ProgressiveMAUVE v.2.4.0 [45] was utilized to compare variations among these 57 species (Table S1). The expansions and contractions of the inverted repeat (IR) regions were examined by analyzing the junctions between the large single-copy (LSC)/IR and small single-copy (SSC)/IR regions, as well as their adjacent genes, using CPJSdraw v0.0.1 [46]. Additionally, the lengths of the LSC, SSC, IR regions, and the total chloroplast genomes for all samples were statistically analyzed in Geneious v10.2.3 (Biomatters Ltd, Auckland, New Zealand). Differences in sequence lengths were assessed using an independent sample *t*-test to evaluate statistical significance for *Polygonatum* and *Heteropolygonatum* by the *t-test* plugin in Rstudio v.2021.09.1.

Phylogenetic analysis

To reconstruct the phylogenetic relationships of *Polygonatum s.l.*, all chloroplast genome sequences involved in this study, including 17 *Heteropolygonatum* and 105 *Polygonatum* sequences, representing seven species of *Heteropolygonatum* and 50 species of *Polygonatum* (Table 1, S1). *Dracaena cambodiana* (MH293451) was chosen as the outgroup. These sequences covered all three major clades within *Polygonatum* (sect. *Verticillata*, sect. *Sibirica*, and sect. *Polygonatum*), and included phylogenetically proximate species to the three medicinal *Polygonatum* taxa. A total of 78 shared protein-coding genes (PCGs) and complete chloroplast genome sequences were extracted and aligned using MAFFT v7.480 in Geneious v10.2.3 with default settings [47]. The phylogenetic relationships of *Polygonatum s.l.* were inferred using both Maximum Likelihood (ML) and Bayesian Inference (BI) methods. The best substitution model was selected using ModelFinder in PhyloSuite v1.2.2 [48]. For the ML analysis, RAxML v2.0.0 was used with the GTRGAMMA model and 1000 bootstrap replicates [49]. BI analysis was performed in MrBayes v3.2 using the Markov Chain Monte Carlo (MCMC) algorithm, running for 20 million generations with trees sampled every 1000 generations [50]. Convergence was assessed by confirming that the average standard deviation of split frequencies was below 0.01. The first 25% of the sampled trees were discarded as burn-in, and the remaining trees were used to construct a consensus tree and estimate posterior probabilities (PP).

Evaluation of species identification based on chloroplast genome

Genetic distance analysis for over two sequences per species followed the methodology outlined by Wang et al. [17]. When the minimum interspecific genetic distance exceeded the maximum intraspecific genetic distance and a barcoding gap was detected, successful species identification was considered achieved [51, 52]. To assess genetic variation within and among the suspicious species, pairwise K2P (Kimura 2-parameter) distances were calculated using MEGA X [53]. Scatter plots were generated to visualize the minimum interspecific K2P distance versus the maximum intraspecific K2P distance for each species, providing a clear illustration of whether the species were accurately defined.

Three molecular species assessment methods, one distance-based and two coalescent-based, were also employed for species delimitation by all available sequences in this study. First, the tree generated using the Bayesian Inference (BI) was converted to NEXUS format and applied to the Bayesian Poisson Tree Processes (bPTP) and Generalized Mixed Yule Coalescent (GMYC) models for automatic species identification using default parameters [32, 34]. Next, the Assemble Species by Automatic Partitioning (ASAP) approach was applied directly to the aligned complete chloroplast genomes for species delimitation with default parameters [33]. Combined with genetic distance analysis, a total of four results for species delimitation were obtained, and the following principles were used to conduct a comprehensive evaluation: (i) if more than or equal to half of the methods supported a particular result, it was considered as the initial assessment; (ii) if the initial assessment aligned with the accepted species classification, it was deemed a trustworthy result; otherwise, the species was considered suspicious. Finally, based on the species delimitation results from each method, this study aimed to evaluate the validity of the currently accepted species delimitations

for *Polygonatum s.l.*, considering whether species were overly subdivided or misidentified.

Identification of species-specific makers for genera and species of Polygonatum s.l

To establish contamination-free authentication of three medicinal significant *Polygonatum* species (*P. cyrtonema*, *P. sibiricum* and *P. kingianum*) against *Heteropolygonatum* adulterants, we engineered an automated screening framework that identifies taxon-specific diagnostic markers through comparative alignment analysis. The pipeline was designed to detect species-specific sites from aligned sequence files, which were categorized as candidate single nucleotide polymorphisms (SNPs) and insertion-deletion (InDel) regions. InDel molecular markers, known for their rich genetic information and cost-effectiveness, are a pivotal tool in species identification [54–56]. For subsequent InDel-specific primer development and validation [57, 58], the following criteria were applied to filter InDel regions: (i) Avoidance of repetitive regions: InDels must not reside within genomic repetitive elements; (ii) Multi-allelic nature: InDels should exhibit multiple allelic states to enhance discriminatory power; (iii) Stability across samples: InDels must exhibit consistent presence/absence patterns across all samples within both *Polygonatum* and its sister genera; (iv) Size threshold: InDel lengths must exceed 50 bp to ensure reliable amplification and detection. For regions that failed to meet the aforementioned criteria, this study prioritized genomic regions harboring overlapping diagnostic loci capable of unambiguously authenticating three medicinal *Polygonatum*. These candidate regions were subsequently validated using the ASAP molecular species delimitation framework.

Development and validation of species-specific sites as molecular barcodes

Primers were designed using Primer Premier v6.0 (http://www.premierbiosoft.com/), with the universal primer *trnH-psbA* (5'-ACTGCCTTGATCCACTTGGC-3' and 5'-CGAAGCTCCATCTACAAATGG-3') for *Polygonatum* included as a control [23, 55]. Although previous studies had validated the effectiveness of chloroplast genome extraction and assembly from the rhizomes of *Polygonatum s.l.* [17], fresh rhizomes were also used in this study for DNA extraction to ensure the universality of this method in practical applications. PCR amplification was performed following a specific protocol described by Cao et al. [59], and 2% agarose gel electrophoresis was conducted for three medicinal *Polygonatum* (*P. cyrtonema*, *P. sibiricum*, and *P. kingianum*) and three *Heteropolygonatum* (representing three main evolutionary clades respectively: *H. alternicirrhosum*, *H. binatifolium*, and *Heteropolygonatum* sp. Kangding). Each species was tested in triplicate to ensure the stability and reproducibility of the results.

Validated regions containing species-specific sites were also subsequently utilized for primer design, serving as species-specific barcodes for robust authentication of medicinal *Polygonatum* materials. PCR amplification and subsequent sequencing protocols were rigorously executed in accordance with the methodology outlined by Dong et al. [60].

Results

General characteristics of newly generated complete Chloroplast genomes within Polygonatum s.l

All 21 newly sequenced complete chloroplast genomes from 13 species of *Polygonatum s.l.* exhibited the typical quadripartite structure. Within the *Polygonatum* genus, genome lengths ranged from 154,578 bp (*P. langyaense*) to 155,987 bp (*P. zanlanscianense*), and within the *Heteropolygonatum* genus, they ranged from 155,278 bp (*H. binatifolium*) to 155,946 bp (*H. alternicirrhosum*; Table 2). The newly obtained lengths of the LSC region lengths in *Polygonatum* ranged from 83,530 to 84,676 bp, while those in *Heteropolygonatum* spanned from 84,324 to 84,974 bp. Similarly, the lengths of the IRa/b regions ranged from 26,295 to 26,414 bp in *Polygonatum* and 26,192 to 26,227 bp in *Heteropolygonatum* (Table 2). Significant intergeneric variation in the lengths of both the LSC and IR regions between the two genera. A *t*-test comparing all sequences in this study revealed significant differences ($P < 0.01$) in the lengths of the LSC and IR regions between *Heteropolygonatum* and *Polygonatum*. In contrast, no significant variation was observed in the SSC regions or other conserved regions of chloroplast genomes ($P > 0.01$; Fig. 2).

Comparative analysis of complete Chloroplast genomes

Each chloroplast genome contained 112 unique genes, including four rRNA genes, 30 tRNA genes, and 78 PCGs. Of these, 15 genes contained a single intron, while *rps12*, *clpP1*, and *ycf3* each included two introns. The IR regions of each chloroplast genome contain 20 genes, consisting of eight PCGs, eight tRNA genes, and four rRNA genes. The average gene density was consistent across all chloroplast genomes, with approximately 0.85 genes per kb and GC content ranging from 37.6 to 37.7% (Table 2). Our comparative analysis of 57 chloroplast genome sequences revealed no structural variations within both *Polygonatum* and *Heteropolygonatum* (Fig. S1). The JLB (LSC/IRb) junction, typically located between the *rps19* and *rpl22* genes was conserved across most plastids in this study, except in *P. sibiricum* and *P. sinopubescens*, which contained 60 bp and 7 bp of *rps19* pseudogenes at the JLB border, respectively (Fig. S2). In contrast, the JSB (IRb/SSC) junction generated *ndhF*

Table 2 Genome information of the newly sequenced samples in this study

Scientific name	GenBank number	The length of chloroplast genome (bp)	The length of IRa/b (bp)	The length of SSC (bp)	The length of LSC (bp)	GC content (%)	Genes/kb
H. alternicirrhosum	PQ591744	155,946	26,227	18,518	84,974	37.6	0.846
H. binatifolium	PQ436974	155,456	26,192	18,518	84,554	37.6	0.849
H. binatifolium	PQ436975	155,463	26,192	18,528	84,551	37.6	0.849
H. binatifolium	PQ591745	155,359	26,213	18,528	84,405	37.6	0.850
H. binatifolium	PQ591746	155,278	26,213	18,528	84,324	37.6	0.850
Heteropolygonatum. sp. Kangding	PQ436973	155,519	26,227	18,518	84,547	37.6	0.849
P. curvistylum	PV055066	155,965	26,413	18,551	84,588	37.7	0.846
P. caulialatum	PV055067	155,335	26,302	18,464	84,267	37.7	0.850
P. cyrtonema var. *gutianshanicuym*	PV055063	155,630	26,370	18,430	84,460	37.7	0.848
P. cyrtonema var. *gutianshanicuym*	PV055053	155,635	26,370	18,432	84,463	37.7	0.848
P. cyrtonema var. *gutianshanicuym*	PV055064	155,630	26,370	18,430	84,460	37.7	0.848
P. franchetii	PV055065	155,931	26,353	18,549	84,676	37.7	0.847
P. jinzhaiense	PV055061	155,515	26,378	18,292	84,467	37.7	0.849
P. jinzhaiense	PV055062	155,520	26,378	18,292	84,472	37.7	0.849
P. langyanese	PV055060	154,578	26,296	18,456	83,530	37.7	0.854
P. macropodum	PV055059	154,613	26,295	18,463	83,560	37.7	0.854
P. nodosum	PV055057	155,211	26,319	18,428	84,145	37.7	0.850
P. nodosum	PV055058	155,205	26,319	18,421	84,146	37.7	0.850
P. punctatum	PV055055	155,274	26,302	18,361	84,309	37.7	0.850
P. punctatum	PV055056	155,329	26,302	18,362	84,363	37.7	0.850
P. zanlanscianense	PV055054	155,987	26,414	18,661	84,498	37.6	0.846

pseudogenes ranging from 22 to 35 bp in most samples, with the exception of *H. ogisui*. The JSA (SSC/IRa) junction between *ycf1* and *trnN-GUU*, and the JLA (IRa/LSC) junction between *rps19* and *trnK-UUU* were found to be highly consistent across all analyzed sequences (Fig. S2).

Phylogenetic analysis

This study reconstructed a phylogenetic framework based on 17 *Heteropolygonatum* and 105 *Polygonatum* sequences, consisting of 57 distinct taxa. The phylogenetic tree topologies, derived from both 78 concatenated PCGs and complete chloroplast genome sequences, were nearly identical. Both *Polygonatum* (BS = 100, PP = 1.00) and *Heteropolygonatum* (BS = 100, PP = 1.00) formed well-supported monophyletic clades, sister to each other (Fig. 1B, S3). Additionally, the inferred phylogeny recovered three sections within *Polygonatum*, including sect. *Verticillata*, sect. *Sibirica*, and sect. *Polygonatum*, with each forming a well-supported monophyletic group (Fig. 1B, S3; BS = 100, PP = 1.00). The newly reported sequences *P. cyrtonema* var. *gutianshanicuym* shared close phylogenetic affinity with the medicinal plant *P. cyrtonema* (BS = 100, PP = 1.00), clustering with *P. jinzhaiensis* in a robust clade (BS = 100, PP = 1.00). In contrast, *P. sibiricum*, another medicinal species, was placed in sect. *Sibirica* as the only member of an independent lineage (Fig. 1B). Despite taxonomic debates continue regarding whether *P. uncinatum* and *P. huanum* should be treated as synonyms of *P. kingianum*, our analysis unambiguously resolved them within the *P. kingianum* clade (BS = 100, PP = 1.00) (Fig. 1B, S3). *Heteropolygonatum* sp. Kangding was resolved into a distinct lineage (Fig. 1B, S3). Furthermore, the analysis revealed that *P. franchetii* (PV055065) differed from previous studies, with these sequences being classified into sect. *Verticillata* and sect. *Polygonatum*, respectively (Fig. 1B, S3).

Molecular species delimitation

In this study, 57 accepted species of *Polygonatum s.l.* were analyzed, of which 30 species were represented by two or more sequences, with no issues in classifying the two genera (*Polygonatum* vs. *Heteropolygonatum*) and three groups (sect. *Polygonatum*, sect. *Sibirica*, and sect. *Verticillata*) across all methods. However, only 47 species were consistently supported by the species delimitation analyses, accounting for 82.46% of the total. Specifically, 64.91% of species were accepted by the ASAP method, 66.67% by PTP, 71.93% by GMYC, and 96.88% by genetic distance (Fig. 3; Table 3). All evaluation analyses failed to confirm independence in these eight *Polygonatum* species (*P. falcatum*, *P. hirtum*, *P. infundiflorum*, *P. multiflorum*, *P. sinopubescens*, *P. urceolatum*, *P. oppositifolium*, and *H. ginfushanicum*). Meanwhile, *P. curvistylum*, *P. cathcartii*, and *H. marmoratum* showed partial support across methods, but did not pass the final evaluation threshold for species recognition (Fig. 3; Table 3).

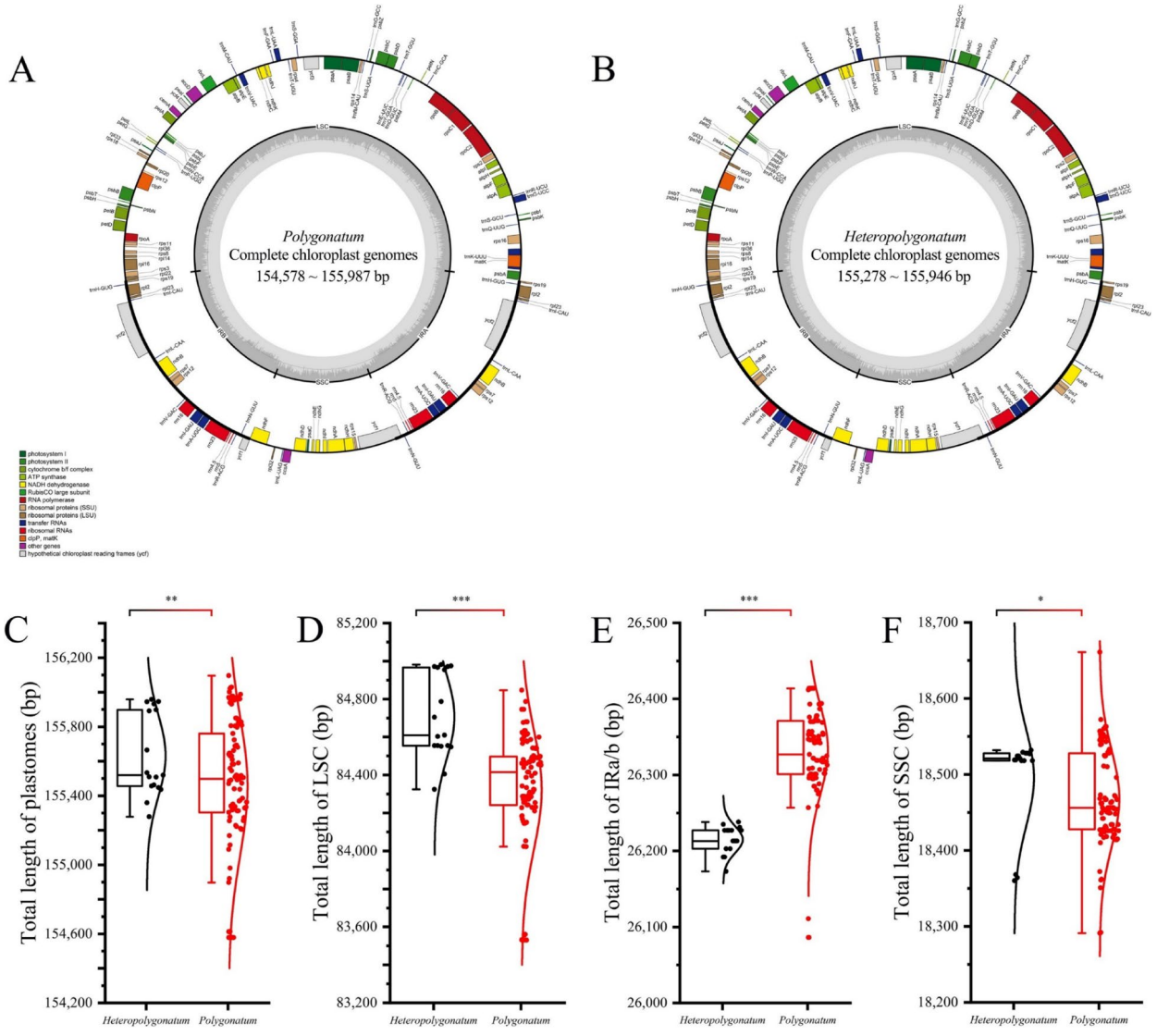

Fig. 2 Complete chloroplast genome map of *Polygonatum* (A) and *Heteropolygonatum* (B) and distribution of lengths and *t*-test results for total chloroplast genomes (C), large single-copy (D), inverted repeat (E), and small single-copy (F) regions of *Polygonatum* and *Heteropolygonatum*. Three asterisks indicate $P < 0.001$; Two asterisks indicate $0.01 > P > 0.001$; one asterisk indicates $P > 0.01$

Identification of species-specific makers for genera and species of Polygonatum s.l

This study identified a total of 563 species-specific sites from *Heteropolygonatum*, 288 from *P. sibiricum*, 259 from *P. kingianum*, and 18 from *P. cyrtonema* through a comparative genomic pipeline (Fig. 4A). Among these, three InDel regions meeting stringent criteria were exclusively identified in *Heteropolygonatum*. Two of these regions exhibited mirror-image symmetry, localized within the inverted repeat (IR) regions of the chloroplast genome, and the longest InDel (156 bp) from three regions was selected for subsequent primer design and validation (Fig. S4). To further distinguish these three traditional medicinal *Polygonatum* species from other species, four candidate regions containing species-specific sites for all three species were screened from the total 565 sites (Fig. 4B). Following the ASAP molecular species delimitation method, only region b (located in the *rps11* gene) consistently demonstrated high discriminatory power across multiple replicates (Fig. 4C). This genomic region exhibits dual functionality: (i) It enables reliable differentiation between these three medicinal *Polygonatum* species and phylogenetically related taxa, and (ii) it provides species-specific diagnostic resolution for unambiguous identification of three individual medicinal *Polygonatum* species (*P. sibiricum*, *P. kingianum*, and *P. cyrtonema*).

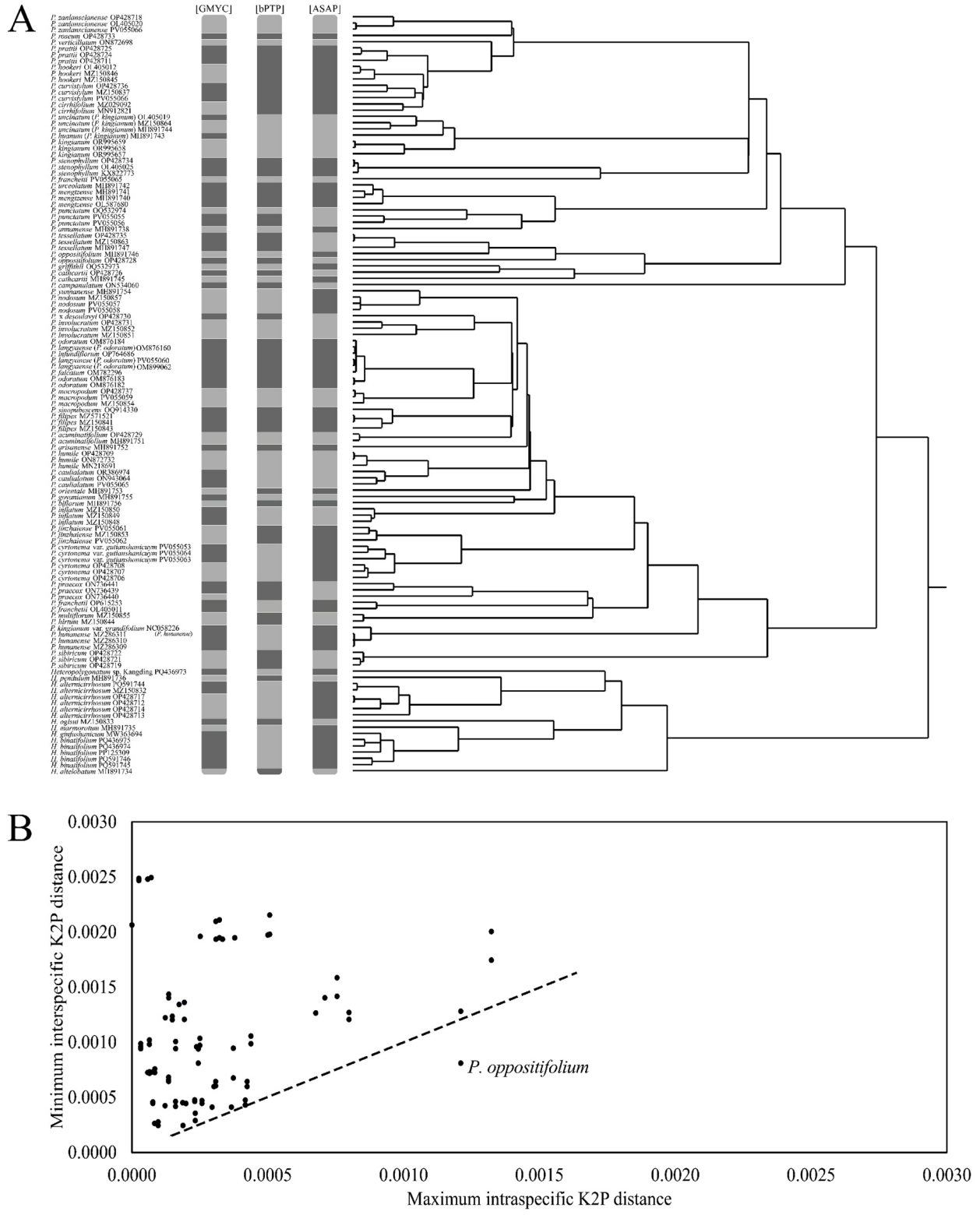

Fig. 3 Results of molecular species assessment. The different blocks in these columns represent the results of the Bayesian Poisson Tree Processes (bPTP), Generalized Mixed Yule Coalescent (GMYC), and Assemble Species by Automatic Partitioning (ASAP), respectively; Genetic distance relationships obtained by the ASAP method are shown on the right (A). Scatter plots of maximum intraspecific Kimura 2-parameter (K2P) distance vs. minimum interspecific K2P distance for complete chloroplast genomes for over two sequences per species (B)

Table 3 Results of multiple molecular species assessments

Species	Sample size	GMYC	PTP	ASAP	Genetic distance	Assessment result
H. altelobatum	1	Yes	Yes	Yes	-	100%
H. ginfushanicum	1	No	No	No	-	0%
H. marmoratum	1	Yes	No	No	-	33%
H. ogisui	1	Yes	Yes	Yes	-	100%
H. pendulum	1	Yes	Yes	Yes	-	100%
Heteropolygonatum sp. Kangding	1	Yes	Yes	Yes	-	100%
P. annamense	1	Yes	Yes	Yes	-	100%
P. arisanense	1	Yes	Yes	Yes	-	100%
P. biflorum	1	Yes	Yes	Yes	-	100%
P. campanulatum	1	Yes	Yes	Yes	-	100%
P. falcatum	1	No	No	No	-	0%
P. govanianum	1	Yes	Yes	Yes	-	100%
P. griffithii	1	Yes	Yes	Yes	-	100%
P. hirtum	1	No	No	No	-	0%
P. huanum (*P. kingianum*)	1	No	Yes	Yes	-	66%
P. infundiflorum	1	No	No	No	-	0%
P. kingianum var. *grandifolium* (*P. hunanense*)	1	Yes	Yes	Yes	-	100%
P. multiflorum	1	No	No	No	-	0%
P. orientale	1	Yes	Yes	Yes	-	100%
P. roseum	1	Yes	Yes	Yes	-	100%
P. siNopubescens	1	No	No	No	-	0%
P. urceolatum	1	No	No	No	-	0%
P. verticillatum	1	Yes	Yes	Yes	-	100%
P. x desoulavyi	1	Yes	Yes	No	-	66%
P. yunnanense (*P. Nodosum*)	1	Yes	Yes	Yes	-	100%
P. acuminatifolium	2	Yes	Yes	Yes	Yes	100%
P. cathcartii	2	No	No	No	Yes	25%
P. cirrhifolium	2	Yes	No	No	Yes	50%
P. oppositifolium	2	No	No	No	No	0%
P. caulialatum	3	Yes	No	No	Yes	50%
P. curvistylum	3	Yes	No	No	Yes	50%
P. cyrtonema	3	Yes	No	No	Yes	50%
P. cyrtonema var. *gutianshanicuym*	3	Yes	No	No	Yes	50%
P. filipes	3	Yes	Yes	Yes	Yes	100%
P. franchetii	3	Yes	Yes	Yes	Yes	100%
P. hookeri	3	Yes	No	No	Yes	50%
P. humile	3	Yes	No	No	Yes	50%
P. hunanense	3	Yes	Yes	Yes	Yes	100%
P. inflatum	3	Yes	Yes	Yes	Yes	100%
P. involucratum	3	Yes	Yes	Yes	Yes	100%
P. jinzhaiense	3	Yes	Yes	No	Yes	75%
P. kingianum	3	Yes	Yes	Yes	Yes	100%
P. langyanese (*P. odoratum*)	3	Yes	Yes	Yes	Yes	100%
P. macropodum	3	Yes	Yes	Yes	Yes	100%
P. mengtzense (*P. punctatum*)	3	Yes	Yes	Yes	Yes	100%
P. Nodosum	3	Yes	Yes	Yes	Yes	100%
P. odoratum	3	Yes	Yes	Yes	Yes	100%
P. praecox	3	No	Yes	Yes	Yes	75%
P. prattii	3	Yes	No	No	Yes	50%
P. punctatum	3	No	No	Yes	Yes	50%
P. sibiricum	3	Yes	Yes	Yes	Yes	100%
P. steNophyllum	3	Yes	Yes	Yes	Yes	100%
P. tessellatum	3	Yes	Yes	Yes	Yes	100%

Table 3 (continued)

Species	Sample size	GMYC	PTP	ASAP	Genetic distance	Assessment result
P. uncinatum (*P. kingianum*)	3	No	Yes	Yes	Yes	75%
P. zanlanscianense	3	Yes	Yes	Yes	Yes	100%
H. binatifolium	5	Yes	Yes	Yes	Yes	100%
H. alternicirrhosum	6	No	Yes	Yes	Yes	75%
Total	**122**	**71.93%**	**66.67%**	**64.91%**	**96.88%**	**82.46%**

* The value of assessment result ≥ 50% were recognized as a confirmed species

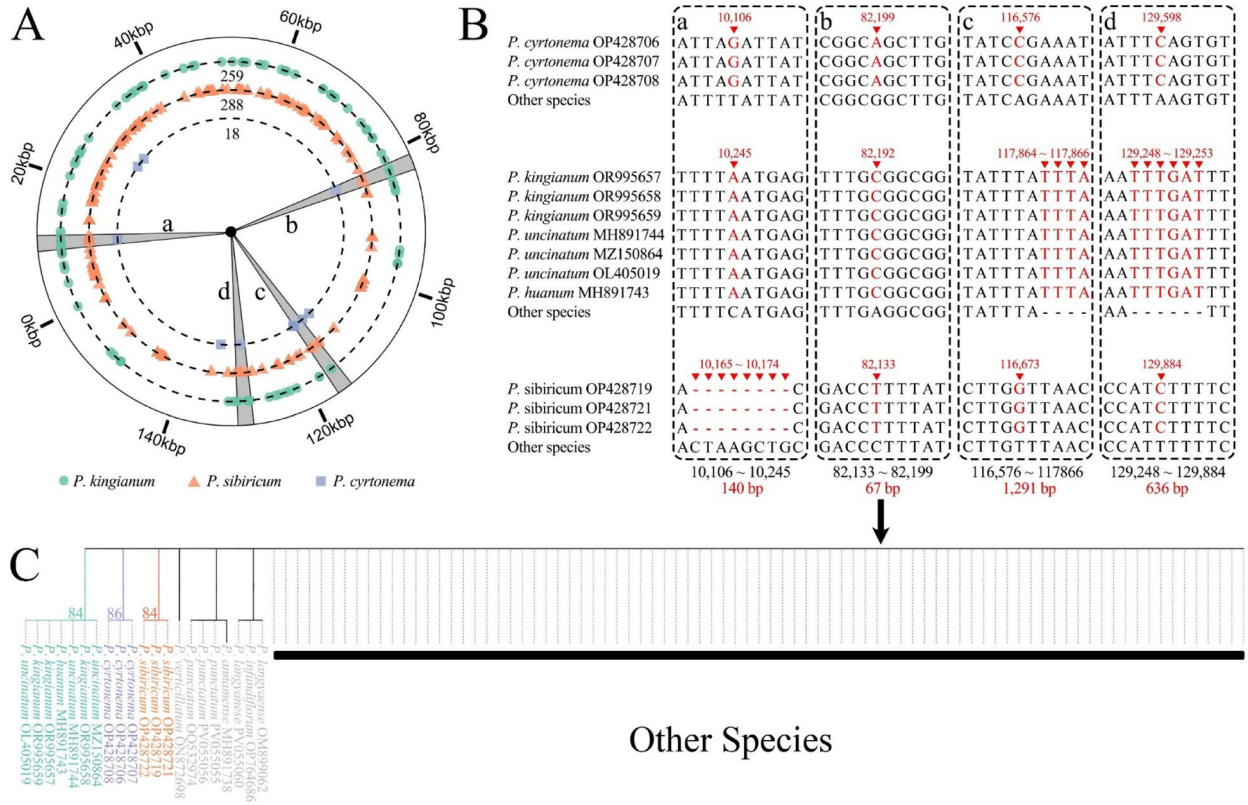

Fig. 4 Identification and validation of species-specific sites for three medicinal *Polygonatum* species. Distribution (A) and categories (B) of species-specific sites; Validation of ASAP molecular species for region b (C)

Development and validation of indels and species-specific primers

To obtain reliable molecular identification markers, InDel primer (5'-GGCAAGAGGATAGCAAGTT-3' and 5'-TGGATTGACTACGGATTCG-3') and SNPs primer (5'-TAATGCTGCGTCTCTTCC-3' and 5'-TGTTACAGATGTACGAGGTC-3') were designed separately for the longest InDel regions and region b. For InDel primer, PCR amplification consistently yielded strong bands, demonstrating the high applicability of the primers (Fig. 5, Fig. S4). Electrophoresis results revealed distinct amplification lengths for *Polygonatum* (ca. 474 bp) and *Heteropolygonatum* (ca. 322 bp), detecting potential contamination of *Heteropolygonatum* and simplifying the process with agarose gel electrophoresis. For the SNP-based primer targeting region b within the *rps11* gene, the marker's specificity and stability were validated across multiple experimental replicates (Fig. 6), further demonstrating its robustness for species-level identification.

Discussion

New insights into taxonomy assessments of Polygonatum s.l. Through Chloroplast genome

Chloroplast genomes have been extensively utilized for species delimitation and phylogenetic reconstruction due to their conserved structure and substantial sequence length [61–64]. In this study, 21 newly sequenced genomes exhibited genomic size, structural organization, and GC content comparable to those reported in previous monocotyledonous plant research, thereby reinforcing the conserved nature of chloroplast genomes across the *Polygonatum s.l.* at a broader taxonomic scale

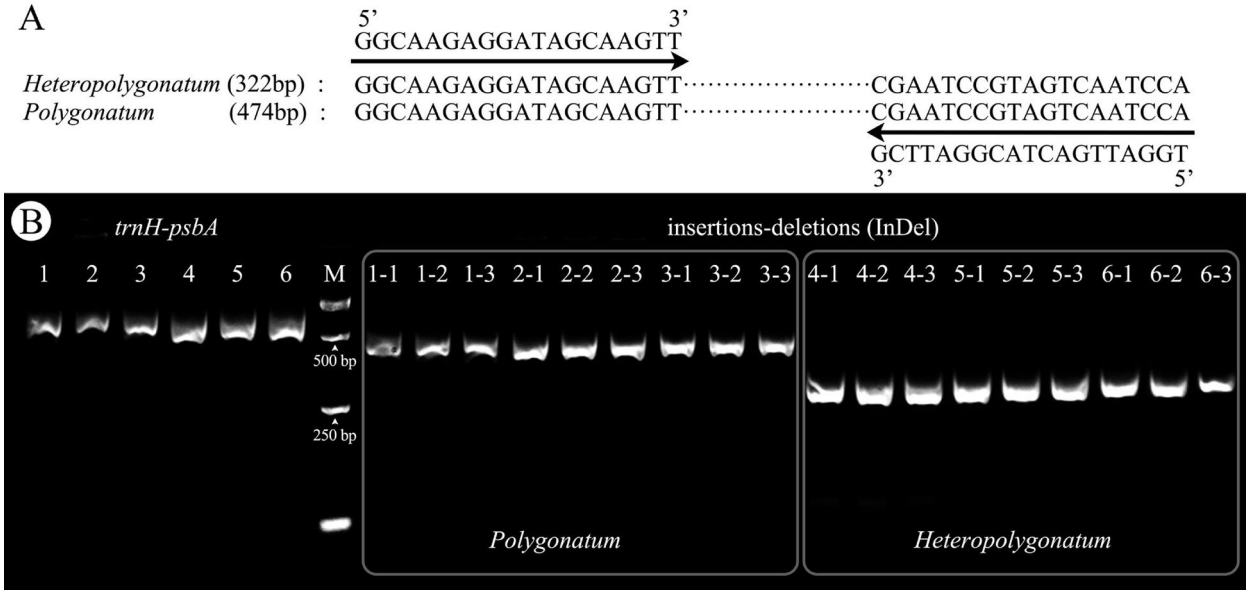

Fig. 5 InDel primer sequences and electrophoretograms for *Polygonatum* and *Heteropolygonatum*. InDel primer sequences, with left and right arrows indicating forward and reverse primers, respectively (A). Electrophoretogram of *trnH-psbA* and InDel sequence fragments in 6 species: *P. cyrtonema* (1), *P. sibiricum* (2), *P. kingianum* (3), *H. alternicirrhosum* (4), *H. ginfushanicum* (5), and *H. binatifolium* (6), each species was tested in triplicate (B)

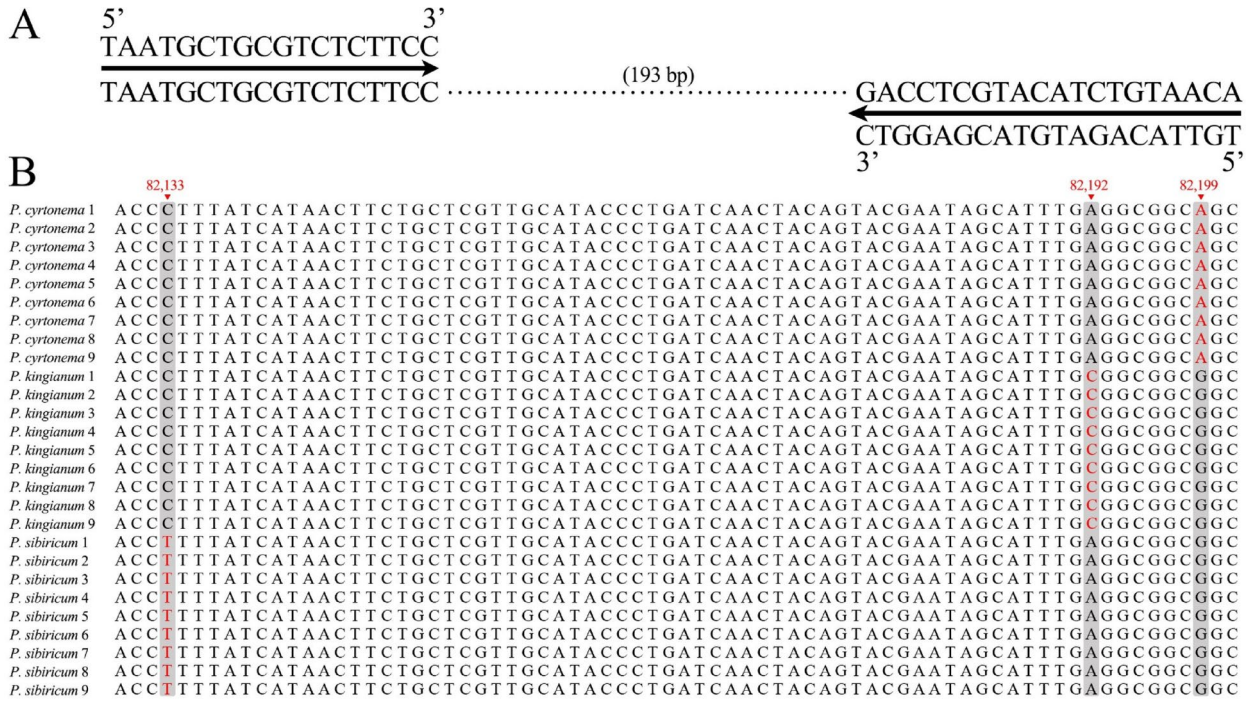

Fig. 6 Species-specific primers and validation results for three medicinal *Polygonatum* species. Species-specific primers sequence (A) and validation results (B)

[65–69]. While earlier investigations primarily on infrageneric lineage classification within *Polygonatum*, few studies have systematically evaluated the efficacy of chloroplast genomes for species identification, especially among medicinally important taxa [2, 70, 71]. To address this gap, we compiled a dataset of 57 species by integrating public databases, representing the most comprehensive plastid-based phylogenetic analysis of *Polygonatum s.l.* to date (Fig. 1; Table S1). Notably, almost all species resolved into distinct, highly supported monophyletic clades (bootstrap values > 99%), validating the three morphological defined infrageneric lineages [2, 24].

Furthermore, multiple molecular species delimitation approaches (ASAP, bPTP, and GMYC) alongside genetic distance thresholds, which improved accuracy and reduced potential misclassification [17]. Three traditional medicinal *Polygonatum* species (*P. sibiricum*, *P. cyrtonema*, and *P. kingianum*) consistently formed distinct clades, even though preliminary analyses showed incongruence (Fig. 3). *P. sibiricu*m maintained its distinct lineage, while *P. cyrtonema* var. *gutianshanicuym*, previously considered a variety of *P. cyrtonema*, was resolved as its closest relative, though both retained independent evolutionary lineages [17]. Interestingly, *P. cyrtonema* and its two phylogenetically proximate species (*P. cyrtonema* var. *gutianshanicuym* and *P. jinzhaiense*) have established three major geographical indication brands. Although no public reports exist of adverse medical events from species mixing, the potential variations in bioactive components—particularly polysaccharides and steroidal saponins—warrant further toxicopharmacological evaluation. Furthermore, integrative taxonomy approaches are recommended to resolve potential oversplitting in these complexes. Such as, the merger of *P. kingianum*, *P. huanum* and *P. uncinatum* was strongly supported in our assessments. These results underscore the power of using comprehensive assessments in resolving conflicts from single-method inferences, and provide a foundation for genetic resource development and conservation prioritization [1]. Furthermore, these findings also support the recent taxonomic revision of *H. binatifolium* as a new combination within *Heteropolygonatum* [21, 72]. The distinct phylogenetic position of, *Heteropolygonatum* sp. Kangding highlights a potentially novel lineage that merits further taxonomic investigation. Such discoveries underscore the value of molecular systematics in biodiversity studies and their application to medicinal resource management.

Novel markers for genera and species identification

Accurate species identification is critical for both conservation and clinical applications of medicinal plants [40]. In *Polygonatum* (including *P. sibiricum*, *P. cyrtonema*, and *P. kingianum*) — three widely used medicinal plants in traditional Chinese medicine — morphological characteristics serve as primary diagnostic markers in their natural state [1]. However, processing for herbal use often erases these traits, especially in tuberous materials. While chloroplast superbarcodes have have proven powerful for species delimitation, their practical implementation faces challenges such as experimental complexity, prolonged processing times, and high costs [17, 73, 74]. Historically, *P. praecox* —a recently described species phenotypically resembling *P. cyrtonema*—was indiscriminately harvested and utilized as the latter in herbal markets [27]. However, integrated evidence from plastid phylogenomics and morphological synapomorphies robustly supports the independent evolutionary status of *P. praecox* [27]. Given its small and restricted population, such misidentification could lead to pharmacovigilance issues and misguided conservation priorities, wasting limited resources.

Traditional barcoding focused on highly variable loci, which often fail in closely related, taxonomically complex genera such as *Polygonatum* [40, 71]. To address these limitations, we designed two new chloroplast-derived molecular markers: (i) InDel primers: genus-level markers for distinguishing *Polygonatum* from phylogenetically proximate genera (e.g., *Heteropolygonatum*), and (ii) species-specific primers: species-level markers targeting three medicinally significant *Polygonatum* species. Compared to prior approaches, our InDel primers allowed rapid genus-level identification (within 2 h) and detection of potential *Heteropolygonatum* contamination using simple gel electrophoresis. Additionally, species-specific primers yielded accurate identification of three medicinal *Polygonatum* species through a single round of Sanger sequencing [39, 70, 75].

Further molecular markers enable rapid screening of wild-collected material, combating illegal harvesting by verifying species identity in confiscated products—a key enforcement tool for protecting overexploited medicinal species. This strategy also offers a model for molecular identification in other medicinal genera, with clear advantages in clinical quality assurance and regulatory frameworks.

Conclusions

We generated 21 complete chloroplast genomes of *Polygonatum s.l.*, expanding current genomic resources to cover over 50% of the taxonomically recognized species in the genus. Phylogenomic analysis resolved previously ambiguous taxa, confirming the distinct evolutionary status of *P. cyrtonema* var. *gutianshanicum* with strong support. The chloroplast genomes showed high discriminatory power, achieving 82.46% identification accuracy across *Polygonatum* s.l. using multiple molecular species delimitation methods. Based on these results, two molecular identification markers were developed: InDel primers targeting a 156-bp indel in the IR region that effectively differentiate *Polygonatum* from *Heteropolygonatum* through gel-based length polymorphism, and species-specific primers in the conserved *rps11* gene, enabling simultaneous identification of three medicinal species (*P. cyrtonema*, *P. kingianum*, and *P. sibiricum*) via a single Sanger sequencing reaction. Furthermore, the findings support prior taxonomic revisions of *P. kingianum* var. *grandifolium* and *H. binatifolium*, and also suggest that *Heteropolygonatum* sp. Kangding represents a distinct lineage that warrants further investigation.

Overall, this integrated approach combining chloroplast genome comparison and marker development provides a robust framework for species authentication and has significant implications for quality control in medicinal applications.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-12012-y.

Supplementary Material 1

Author contributions

Y. Hu and J. Shao wrote the main manuscript text. S. Wang, Z. Xu, S. Yan, and M. Ali contributed to data validation, investigation, and formal analysis. Z. Li contributed to conceptualization, data curation, and formal analysis. M. Ali also participated in writing – review & editing. J. Shao was responsible for project administration, validation, supervision, and funding acquisition. All authors reviewed the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (Nos. 32070370 and 32470380).

Data availability

All newly annotated complete chloroplast genome sequences in this study are available from the Genbank with accession numbers: PV055053.to PV055067, PQ436973 to PQ436975 and PQ591744 to PQ591746.

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]College of Life Sciences, Anhui Normal University, Wuhu 241000, Anhui, China
[2]School of Ecology and Environment, Anhui Normal University, Wuhu 241000, Anhui, China
[3]The Anhui Provincial Key Laboratory of Biodiversity Conservation and Ecological Security in the Yangtze River Basin, Wuhu 241000, Anhui, China
[4]Center for Integrative Conservation & Yunnan Key Laboratory for Conservation of Tropical Rainforests and Asian Elephants, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, Yunnan, China

Received: 4 May 2025 / Accepted: 18 August 2025
Published online: 02 September 2025

References

1. Chen XQ, Tamura MN. Polygonatum. In: Wu ZY, Peter PH, editors. Flora of China. Beijing: Science; St. Louis: Missouri Botanical Garden Press;; 2000. pp. 223–32.
2. Floden AJ, Schilling EE. Using phylogenomics to reconstruct phylogenetic relationships within tribe polygonateae (Asparagaceae), with a special focus on *Polygonatum*. Mol Phylogenet Evol. 2018;129:202–13.
3. Jeffrey C. The genus *Polygonatum* (Liliaceae) in Eastern Asia. Kew Bull. 1980;34:435–71.
4. Jeffrey C. Further note on Eastern Asian *Polygonatum* (Liliaceae). Kew Bull. 1982;37:335–9.
5. Tamura MN, Ogisu M, Xu J. *Heteropolygonatum*, a new genus of the tribe polygonateae (Convallariaceae) from West China. Kew Bull. 1997;52:949–56.
6. Shan F, Huang LQ, Guo J, Chen M. History and development of one root of medicine and food. Chin Bull Life Sci. 2015;27:1061–9.
7. Shi Y, Liu JJ, Si D, Golding JB, Pristijono P, Li YX, He FL, Zhang XF, Han ZG, Wu LS, Chen DH, Si JP. Huangjing—From medicine to healthy food and diet. Food Front. 2023;4:1068–90.
8. Li XL, Ma RH, Ni ZJ, Thakur K, Cespedes-Acuna CL, Wang S, Zhang JG, Wei ZJ. Dioscin inhibits human endometrial carcinoma proliferation via G0/G1 cell cycle arrest and mitochondrial–dependent signaling pathway. Food Chem Toxicol. 2021;148:111941.
9. Shen WD, Li XY, Deng YY, Zha XQ, Pan LH, Li QM, Luo JP. *Polygonatum Cyrtonema* Hua polysaccharide exhibits anti–fatigue activity via regulating osteocalcin signaling. Int J Biol Macromol. 2021;175:235–41.
10. He LL, Yan BX, Yao CY, Chen XY, Li LW, Wu YJ, Song ZJ, Song SS, Zhang ZF, Luo P. Oligosaccharides from *Polygonatum Cyrtonema* hua: structural characterization and treatment of LPS–induced peritonitis in mice. Carbohydr Polym. 2021;255:117392.
11. Huang S, Yuan HY, Li WQ, Liu XY, Zhang XJ, Xiang DX, Luo SL. Polysaccharides protect against MPP–induced neurotoxicity via the akt/mtor and Nrf2 pathways. Oxid Med Cell Longev. 2021;8843899.
12. Si JP, Zhu YX. Polygonati rhizome–A new high–quality crop with great potential and not occupying farmland. Sci Sin Vitae. 2021;51:1477–84.
13. De Queiroz K. Species concepts and species delimitation. Syst Biol. 2007;56:879–86.
14. Liu JQ. The integrative species concept and species on the speciation way. Biodiv Sci. 2016;24:1004–8.
15. Hong DY. Gen–morph species concept-A new and integrative species concept for outbreeding organisms. J Syst Evol. 2020;58:725–42.
16. Chen SL, Pang XH, Song JY, Shi LC, Yao H, Han JP, Leon C. A renaissance in herbal medicine identification: from morphology to DNA. Biotechnol Adv. 2014;32:1237–44.
17. Wang SY, Zhou N, Shi NX, Zhang GF, Liu HY, Guo XR, Ji YH. Testing and using complete plastomes for authentication of medicinal *Polygonatum* species (Asparagaceae). Ind Crops Prod. 2023;197:116557.
18. Tamura MN, Chen X, Turland NJ. A new combination in heteropolygonatum (Convallariaceae, Polygonateae). Novon. 2000;2:156–7.
19. Floden AJ. A new combination in *Polygonatum* (Asparagaceae) and the reinstatement of *P. mengtzense*. Ann Botanici Fennici. 2014;51:106–16.
20. Floden AJ. New names in *Heteropolygonatum*. (Asparagaceae) Phytotaxa. 2014;188:218–26.
21. Zhu HX, Huang FY, Cheng HY, Zhang HJ, Dai QL, Yi SR. *Polygonatum binatifolium*, a new species of *Polygonatum* (Asparagaceae) from guizhou, China. Phytotaxa. 2022;549:247–52.
22. Zhou YL, Jiang H, Hu GW, Wang QF. *Heteropolygonatum binatifolium*, a new combination in *Heteropolygonatum* (Asparagaceae, Polygonateae). Phytotaxa 2025;682:274–280.
23. Zhao LH, Zhou SD, He XJ. A phylogenetic study of Chinese *Polygonatum* (Polygonateae, Asparagaceae). Nord J Bot. 2019;e02019.
24. 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.
25. 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.
26. Xia M, Liu Y, Liu J, Chen D, Shi Y, Chen Z, Chen D, Jin R, Chen H, Zhu S. Out of the Himalaya–Hengduan mountains: phylogenomics, biogeography and diversification of *Polygonatum* mill. (Asparagaceae) in the Northern hemisphere. Mol Phylogenet Evol. 2022;169:107431.
27. Hu YF, Liu YJ, Ali M, Wu W, Li XH, Cheng LS, Shao JW. *Polygonatum praecox* (Asparagaceae), a new species from mid–eastern China revealed by morphological and molecular evidence. PhytoKeys. 2022;211:125–38.
28. Zhang D, Ren J, Jiang H, Wanga VO, Dong X, Hu G. Comparative and phylogenetic analysis of the complete Chloroplast genomes of six *Polygonatum* species (Asparagaceae). Sci Rep. 2023;13:7237.
29. Guo XR, Shi NX, Xie PX, Zhang GF, Liu HY, Ji YH. Plastome sequencing for accurate and effective authentication of *Polygonatum Kingianum* (Asparagaceae). Ind Crops Prod. 2022;184:115056.

30. Yan MX, Dong SJ, Gong QY, Xu Q, Ge YQ. Comparative Chloroplast genome analysis of four *Polygonatum* species insights into DNA barcoding, evolution, and phylogeny. Sci Rep. 2023;13:16495.
31. Luo AR, Ling C, Ho SYW, Zhu CD. Comparison of methods for molecular species delimitation across a range of speciation scenarios. Syst Biol. 2018;67:830–46.
32. Fujisawa T, Barraclough TG. Delimiting species using single–locus data and the generalized mixed Yule coalescent approach: a revised method and evaluation on simulated data sets. Syst Biol. 2013;62:707–24.
33. Puillandre N, Brouillet S, Achaz G. ASAP: assemble species by automatic partitioning. Mol Ecol Resour. 2021;21:609–20.
34. Zhang J, Kapli P, Pavlidis P, Stamatakis A. A general species delimitation method with applications to phylogenetic placements. Bioinformatics. 2013;29:2869–76.
35. Dong W, Liu Y, Li E, Xu C, Sun J, Li W, Zhou S, Zhang Z, Suo Z. Phylogenomics and biogeography of *Catalpa* (Bignoniaceae) reveal incomplete lineage sorting and three dispersal events. Mol Phylogenet Evol. 2022;166:107330.
36. Bennett EP, Petersen BL, Johansen IE, Niu Y, Yang Z, Chamberlain CA, Met O, Wandall HH, Frodin M. INDEL detection, the 'achilles heel' of precise genome editing: a survey of methods for accurate profiling of gene editing induced indels. Nucleic Acids Res. 2020;48:11958–81.
37. Seo JH, Dhungana SK, Kang BK, Baek IY, Sung JS, Ko JY, Jung CS, Kim KS, Jun TH. Development and validation of SNP and indel markers for Pod–Shattering tolerance in soybean. Int J Mol Sci. 2022;23:2382.
38. Shivaprasad KM, Aski M, Mishra GP, Sinha SK, Gupta S, Mishra DC, Singh AK, Singh A, Tripathi K, Kumar RR, Kumar A, Kumar S, Dikshit HK. Genome–wide discovery of indels and validation of PCR-Based indel markers for earliness in a RIL population and genotypes of lentil (Lens culinaris Medik). PLoS ONE. 2024;19:e0302870.
39. Kim Y, Choi H, Shin J, Jo A, Lee KE, Cho SS, Hwang YP, Choi C. Molecular discrimination of *Cynanchum wilfordii* and *Cynanchum auriculatum* by indel markers of Chloroplast DNA. Molecules. 2018;23:1337.
40. Wang S, Wang YH, Sun JH, Cui XY, Li E, Wang RS, Li Q, Zhang PF, Dong WP, Guo LP, Huang LQ. Comparative Chloroplast genome analyses provide new insights into molecular markers for distinguishing arnebiae radix and its substitutes (tribe lithospermeae, Boraginaceae). Phytomedicine. 2025;136:156338.
41. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:884–90.
42. Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. GetOrganelle: a fast and versatile toolkit for accurate de Novo assembly of organelle genomes. Genome Biol. 2020;21:241.
43. Qu XJ, Moore MJ, Li DZ, Yi TS. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods. 2019;15:50.
44. Greiner S, Lehwark P, Bock R. Organellar genome DRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47:59–64.
45. Darling AE, Mau B, Perna NT. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE. 2010;5:e11147.
46. Li H, Guo Q, Xu L, Gao H, Liu L, Zhou X. CPJSdraw: analysis and visualization of junction sites of Chloroplast genomes. PeerJ. 2023;11:e15326.
47. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
48. Zhang D, Gao F, Jakovlić 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.
49. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post–analysis of large phylogenies. Bioinformatics. 2014;30:1312–3.
50. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBaYes 3.2: efficient baYesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61:539–42.
51. Collins R, Cruickshank R. The seven deadly sins of DNA barcoding. Mol Ecol Resour. 2013;13:969–75.
52. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proc Biol Sci. 2003;270:313–21.
53. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35:1547–9.
54. Lv Y, Liu Y, Zhao H. mInDel: a high-throughput and efficient pipeline for genome–wide indel marker development. BMC Genomics. 2016;17:290.
55. Hamilton MB. Four primer pairs for the amplification of Chloroplast noncoding regions with intraspecific variation. Mol Ecol. 1999;8:521–3.
56. Zhao LH, Zhou SD, He XJ. A phylogenetic study of Chinese *Polygonatum* (Polygonateae, Asparagaceae). Nord J Bot. 2019;2019:e02019.
57. Jain A, Roorkiwal M, Kale S, Garg V, Yadala R, Varshney RK. InDel markers: an extended marker resource for molecular breeding in Chickpea. PLoS ONE. 2019;14:e0213999.
58. Wu XY, Xu P, Wu XH, Wang BG, Lu ZF, Li GJ. Development of insertion and deletion markers for bottle gourd based on restriction Site-associated DNA sequencing data. Hortic Plant J. 2017;3:13–6.
59. Cao DL, Zhang XJ, Xie SQ, Fan SJ, Qu XJ. Application of Chloroplast genome in the identification of traditional Chinese medicine viola Philippica. BMC Genomics. 2022;23:540.
60. Dong W, Liu H, Xu C, Zuo Y, Chen Z, Zhou S. A Chloroplast genomic strategy for designing taxon specific DNA mini–barcodes: a case study on ginsengs. BMC Genet. 2014;15:138.
61. Ahmed I, Bigg PJ, Matthews PJ, Collins LJ, Hendy MD, Lockhart PJ. Mutational dynamics of *Aroid* Chloroplast genomes. Genom Biol Evol. 2012;4:1316–23.
62. Henriquez CL, Abdullah, Ahmed I, Carlsen MM, McKain MR. Evolutionary dynamics of Chloroplast genomes in subfamily *Aroideae* (Araceae). Genomics. 2020;112:2349–60.
63. Shahzadi I, Abdullah, Mehmood F, Ali Z, Ahmed I, Mirza B. Chloroplast genome sequences of *Artemisia maritima* and *Artemisia absinthium*: comparative analyses, mutational hotspots in genus *Artemisia* and phylogeny in family Asteraceae. Genomics. 2020;112:1454–63.
64. Yu X, Zuo L, Lu D, Lu B, Yang M, Wang J. Comparative analysis of Chloroplast genomes of five *Robinia* species: genome comparative and evolution analysis. Gene. 2019;689:141–51.
65. Cui G, Wang C, Wei X, Wang H, Wang X, Zhu X, Li J, Yang H, Duan H. Complete Chloroplast genome of *Hordeum brevisubulatum*: genome organization, synonymous codon usage, phylogenetic relationships, and comparative structure analysis. PLoS ONE. 2021;16:e0261196.
66. Freudenberg J, Wang M, Yang Y, Li W. Partial correlation analysis indicates causal relationships between GC–content, exon density and recombination rate in the human genome. BMC Bioinformatics. 2009;10:66.
67. Liang H, Zhang Y, Deng J, Gao G, Ding C, Zhang L, Yang R. The complete Chloroplast genome sequences of 14 *Curcuma* species: insights into genome evolution and phylogenetic relationships within zingiberales. Front Genet. 2020;11:802.
68. Lin RC, Ferreira BT, Yuan YW. The molecular basis of phenotypic evolution: beyond the usual suspects. Trends Genet. 2024;40:668–80.
69. Zhang L, Kasif S, Cantor CR, Broude NE. GC/AT-content spikes as genomic Punctuation marks. Proc Natl Acad Sci U S A. 2004;101:16855–60.
70. Wang W, Mauleon R, Hu Z, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature. 2018;557:43–9.
71. Wang J, Qian J, Jiang Y, Chen X, Zheng B, Chen S, Yang F, Xu Z, Duan B. Comparative analysis of Chloroplast genome and new insights into phylogenetic relationships of *Polygonatum* and tribe polygonateae. Front Plant Sci. 2022;13:882189.
72. Zhou YL, Jiang H, Hu GW, Wang QF. *Heteropolygonatum binatifolium*, a new combination in *Heteropolygonatum* (Asparagaceae, Polygonateae). Phytotaxa. 2025;682:274–80.
73. Hollingsworth PM, Li DZ, Michelle VDB, Twyford AD. Telling plant species apart with DNA: from barcodes to genomes. Philos Trans R Soc B Biol Sci. 2016;371:20150338.
74. Kane NS, Sveinsson S, Dempewolf H, Yang JY, Zhang D, Engels JMM, Cronk Q. Ultra-barcoding in Cacao (Theobroma spp.; Malvaceae) using whole Chloroplast genomes and nuclear ribosomal DNA. Am J Bot. 2012;99:320–9.
75. Tian X, Guo J, Zhou X, Ma K, Ma Y, Shi T, Shi Y. Comparative and evolutionary analyses on the complete plastomes of five *Kalanchoe* horticultural plants. Front Plant Sci. 2021;12:705874.

Publisher's note