BMC Plant Biology

RESEARCH Open Access

Comparative chloroplast genome of six species in Hypoxidaceae from China: insights into phylogenetic relationships and molecular marker development

Dong Ma[1,2†], Qin Tian[1,3†], Yunqiang Wang[4], Hanning Duan[1,5], Yuan Zhang[1], Yan Luo[6*] and Lu Li[1,2*]

Abstract

Background The family Hypoxidaceae (order Asparagales) is a predominantly Southern Hemisphere lineage comprising approximately 11 genera and 200 species, many of which possess significant medicinal and ornamental value. Despite their economic importance, Hypoxidaceae has received limited research attention, leading to problematic identification of species and misuse of wild resources in traditional medicine markets. Taxonomically, the phylogenetic position of Hypoxidaceae and the intergeneric relationships within this family remain controversial and unresolved, particularly concerning the delimitation of *Curculigo* and *Molineria*. Previous studies based on morphological traits and molecular markers have yielded inconsistent results, highlighting the need for more robust genomic evidence. In angiosperms, complete chloroplast genomes have proven highly effective in resolving systematic uncertainties considering their conserved structure and high informational content. However, such genomic data remain scarce for Hypoxidaceae, limiting phylogenetic clarity. In this research, the complete chloroplast genomes of six species representing three key genera (*Curculigo*, *Molineria*, and *Hypoxis*) were sequenced and characterized for a comparative and phylogenetic analysis.

Results The chloroplast genomes of six species exhibited conserved quadripartite structures, measured 157,472 bp to 158,550 bp in length. The overall GC content of these genomes ranged between 37.3 and 37.5%. Gene annotations identified 132 genes, 19 duplicated in the inverted repeat regions, and had complete *ndh* gene. Comparative analysis of six complete chloroplast genomes revealed highly similarity, but they were varied in repeats sequence, codon usage bias, contractions and expansions in the IR region. Five molecular markers showed the highest degree of variability between the six cp genomes. Phylgenetic analysis based on cp genomic data confirmed that Hypoxidaceae was a monophyly, being a sister to Asteliaceae with higher supports than the previous research. Three main clades

†Dong Ma and Qin Tian contributed equally to this work and should be considered as co–first authors.

*Correspondence:
Yan Luo
luoyan@xtbg.org.cn
Lu Li
lilu@swfu.edu.cn

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

were recognized in Hypoxidaceae, including *Curculigo* clade, *Hypoxis* clade, and *Pauridia–Empodium* clade. And what's more, *Curculigo* clade could be divided into three subclades, containing *Molineria* subclade, *Curculigo* subclade, and Seychellean subclade, indicating significantly phylogenetic insights.

Conclusions The complete cp genomes of six species of three representative genera from Hypoxidaceae were sequenced and analyzed in detail, including the general data on the genome length, repeat sequence, codon usage, IR expansion and contraction, structural comparison and divergence hotspot identification analyses, and phylogenetic analysis. A comparative analysis revealed that the cp genome was highly consistent of four *Molineria* species, but varied greatly at the generic level between *Hypoxis, Curculigo*, and *Molineria*, which could be used for generic delimitation. Five DNA barcodes (*psbK–psbI*, *rpoB–trnC, ndhF–rpl32, ycf1,* and *trnE–trnT*) were selected for authentication of Hypoxidaceae medicinal materials. Hypoxidaceae was a monophyletic lineage, containing three major clades, being a sister to Asteliaceae with stronger supports than before. The three main clades in Hypboxidaceae were re-confirmed as the three stable lineages for this family. In the *Curculigo* Clade, three subclades were identified with significant phylogenetic insights. The phylogenetic evidence presented here, combined with distinct chloroplast genome features, supports *Molineria* Subclade separated from *Curuculigo* Subclade, being a monophyletic group by transferring *Sinocurculigo taishanica* and two Borneo *Curculigo* species into *Molineria*. Further research should provide a better understanding of the intergeneric relationships among Hypoxidaceae, adding more genomic data with extensive samplings across the center distribution of Southern Hemisphere.

Keywords Chloroplast genome, DNA barcode, *Curculigo*, *Molineria*, Systematic relationship

Background

The monocot family Hypoxidaceae R. Br. consists of 11 genera and nearly 200 species, primarily distributed across the Southern Hemisphere, with only a limited number extending into the Northern Hemisphere [1, 2]. In China, four genera — *Hypoxis* L., *Curculigo* Gaertn., *Molineria* Colla, and *Sinocurculigo* Z. J. Liu, L. J. Chen & K. Wei Liu — are documented, encompassing a total of nine species [2–4]. These include one species from *Hypoxis* (*H. aurea* Lour.), two from *Curculigo* (*C. orchioides* Gaertn. and *C. glabrescens* (Ridl.) Merr.), five from *Molineria* (*M. capitulate* (Lour.) Herb., *M. breviscapa* S. C. Chen, *M. crassifolia* Baker, *M. gracilis* Kurz, and *M. sinensis* S. C. Chen), and one from the newly described genus *Sinocurculigo* (*S. taishanica* Z. J. Liu, L. J. Chen & K. Wei Liu) [2–4]. Members of Hypoxidaceae are herbaceous plants, typically under 20 cm in height, though some species may exceed one meter. All possess underground storage organs, including tuberous or elongated rhizomes or corms [1, 3, 4]. Their floral morphology follows the classic lily-like monocot pattern, characterized by three sepals, three petals, six stamens, and a trimerous gynoecium [1, 5]. Several species are valued in traditional Chinese medicine, most notably *Curculigo orchioides* (known as 'Xianmao') [6], along with *Molineria capitulata*, *M. breviscapa*, and *Hypoxis aurea* [7–9]. However, due to limited taxonomic and pharmacological research, wild Hypoxidaceae resources in East Asian markets are frequently misidentified or misused [9–11], highlighting the need for further research.

The phylogenetic position of Hypoxidaceae has historically been uncertain. Initially, its members were classified as the tribe Hypoxideae within Amaryllidaceae J. St.-Hil. [12], a treatment later followed by other authors [13], including Chinese taxonomists [3, 4]. However, Hypoxideae differs from the rest of Amaryllidaceae in lacking bulbs and umbellate inflorescences, leading to its eventual elevation to family rank as Hypoxidaceae [14]. Early morphological studies suggested a sister relationship between Hypoxidaceae and Orchidaceae Juss., as both families share epigynous flowers, absence of septal nectaries, and successive microsporogenesis [15–20]. Some researchers even proposed that apostasioid orchids (Apostasioideae Rchb.) might have derived from Hypoxidaceae, given their strong morphological similarities — particularly with *Curculigo* and *Molineria* [15, 16, 18]. In contrast, molecular phylogenetic analysis places Hypoxidaceae within a highly supported Astelioid clade, which also includes Asteliaceae Dumort., Blandfordiaceae R. Dahlgren & Clifford, Lanariaceae H. Huber ex R. Dahlgren, and Boryaceae M. W. Chase, Rudall & Conran, which are part of the 'lower Asparagales' [19, 21–25]. Meanwhile, Amaryllidaceae clusters within the 'higher asparagoid', whereas Orchidaceae occupies a basal position relative to other Asparagales [19, 22, 23, 26, 27].

The intergeneric delimitations within Hypoxidaceae remain unresolved and become challenging despite extensive study [2, 19, 21, 24, 25, 28–30]. Although molecular phylogenetic analyses have consistently identified three major clades, these lineages lack clear morphological or anatomical synapomorphies, resulting in a persistent discordance between molecular and phenotypic data [2, 21, 24, 31]. A particularly contentious issue involves the delimitation of *Curculigo* and *Molineria*, which remains poorly supported due to overlapping and ambiguous morphological traits. Key diagnostic

characters — such as rhizome structure, presence/absence of a fruit beak, and anther symmetry—have proven unreliable, leading to frequent movement of species between the two genera [19, 25, 28, 30]. Molecular studies further complicate the matter, as *Curculigo* and *Molineria* were grouped into a broadly defined *Curculigo* clade [21, 24]. Surprisingly, despite molecular evidence suggested a merged *Curculigo-Molineria* clade, two new monotypic genera—*Sinocurculigo* and *Neofriedmannia* —were recently proposed based on morphological distinctions. *Sinocurculigo* was segregated due to its unique stigma, ovary, and seed morphology, despite clustering within the *Molineria-Curculigo* subclade in plastid (*rbc*L and *mat*K) phylogeny [2]. Similarly, *Curculigo seychellensis* was proposed into a monotypic genus of *Neofriedmannia* based on its exceptionally long fruit beak (~ 12 cm) and phylogenetic placement within a *Seychellean* subclade alongside two endemic *Hypoxidia* species [31, 32]. The recognition of these monotypic genera underscores the ongoing taxonomic tension surrounding *Curculigo* and *Molineria*, highlighting the need for integrative studies combining morphology, molecular phylogenomics, and biogeography to resolve these persistent taxonomic issues.

The chloroplast (cp) genome has become a pivotal tool in angiosperm phylogenetics due to its conserved structure, low recombination rate, and predominantly uniparental inheritance [5, 22]. Its utility is well-demonstrated in resolving long-standing phylogenetic and intergeneric delimitation issues across diverse angiosperm families, including Lauraceae Juss. [33], Ranunculaceae Juss. [34], Orchidaceae [35, 36], and Amaryllidaceae [37]. Furthermore, cp genomic data have effectively clarified species boundaries in taxonomically complex genera of medicinal importance, such as *Artemisia* L. [38], *Panax* L. [39], *Physalis* L. [40], *Paeonia* L. [41], and *Salvia* L. [42]. Despite its proven efficacy, cp genomic resources for Hypoxidaceae remain strikingly limited, with only a single published cp genome—*Curculigo orchioides*—currently available [43]. Therefore, we sequenced, characterized, and compared the complete chloroplast genomes of six species, representing three key genera within Hypoxidaceae from China, including *Hypoxis aurea*, *Curculigo orchioides*, and four *Molineria* species (*M. capitulata*, *M. crassifolia*, *M. sinensis*, and *M. breviscapa*). This study has two primary objectives: to annotate and compare the structural features of the six cp genomes, identifying conserved and variable regions; to reconstruct robust phylogenetic trees, providing new genomic insights into the evolutionary relationships within Hypoxidaceae, with particular emphasis on the contentious delimitation between *Curculigo* and *Molineria*.

Results

General features of the chloroplast genome of Hypoxidaceae

The lengthes of six complete cp genomes were varied from 157,472 bp (*Curculigo orchioides*) to 158,550 bp (*Molineria capitulata*), with a typical quadripartite structure, holding a large single–copy (LSC) region and a small single–copy (SSC) region separated by two IR regions (Fig. 1 and Table S1). The GC content of six cp genomes was highly conserved, with overall GC content (37.3 to 37.5%) and AT–rich (Table S1). IR regions showed a higher GC content (42.5 to 42.6%) than LSC region (35.3 to 35.5%) and SSC region (31.2 to 31.4%) (Table S1).

Six cp genomes shared 132 gene numbers, including 86 proteins–coding genes, eight rRNAs, and 38 tRNA–coding genes (Fig. 1 and Table S1). The 11 *ndh* (NA(D)H dehydrogenase) genes (*ndh A–K*) were identified in the six cp genomes (Table S2). Among these genes, there were a total of 19 double–copy genes, including eight tRNA–encoding genes (*trnA–UGC*, *trnH–GUG*, *trnI–CAU*, *trnI–GAU*, *trnL–CAA*, *trnN–GUU*, *trnR–ACG*, and *trnV–GAC*), four rRNA encoding genes (*rrn16*, *rrn23*, *rrn4.5*, and *rrn5*), and seven proteins–encoding genes (*ndhB*, *rpl2*, *rpl23*, *rps12*, *rps19*, *rps7*, and *ycf2*) (Table S2). Two introns were detected in each of the three protein–coding genes, *rps12*, *clpP1*, and *ycf3*, while a single intron was respectively detected in each of nine proteins–coding genes (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, *rpoC1*, and *rps16*) and six tRNA genes (*trnA–UGC*, *trnG–UCC*, *trnI–GAU*, *trnK–UUU*, *trnL–UAA*, and *trnV–UAC*) (Table S2).

Repeat sequence analysis

The MISA online analysis software was utilized to statistically analyze the nucleotide repeat type and number of SSRs (simple sequence repeats) in six cp genomes from Hypoxidaceae. Among these, *Molineria breviscapa* exhibited the largest number of SSR with 92, while *Hypoxis aurea* had the smallest at 82 (Table S3). Six types of SSRs were screened in six species. Four types of SSRs were consistently identified in six cp genomes, including mononucleotide (P1), dinucleotide (P2), trinucleotide (P3), and tetranucleotide (P4) (Fig. 2, Table S3). Pentanucleotide (P5) was present in the five species, but absent in *M. crassifolia*. Conversely, hexanucleotides (P6) was only found in *M. crassifolia* (Fig. 2, Table S3). The SSRs identified in cp genome of six species from Hypoxidaceae were primarily composed of A and T nucleotides, with over half of the SSR sequences consisting solely of A or T bases (Table S3). It was revealed that the majority of SSR were distributed in the intergenic regions, while the number of SSR in the intron and gene regions was similar (Fig. 2).

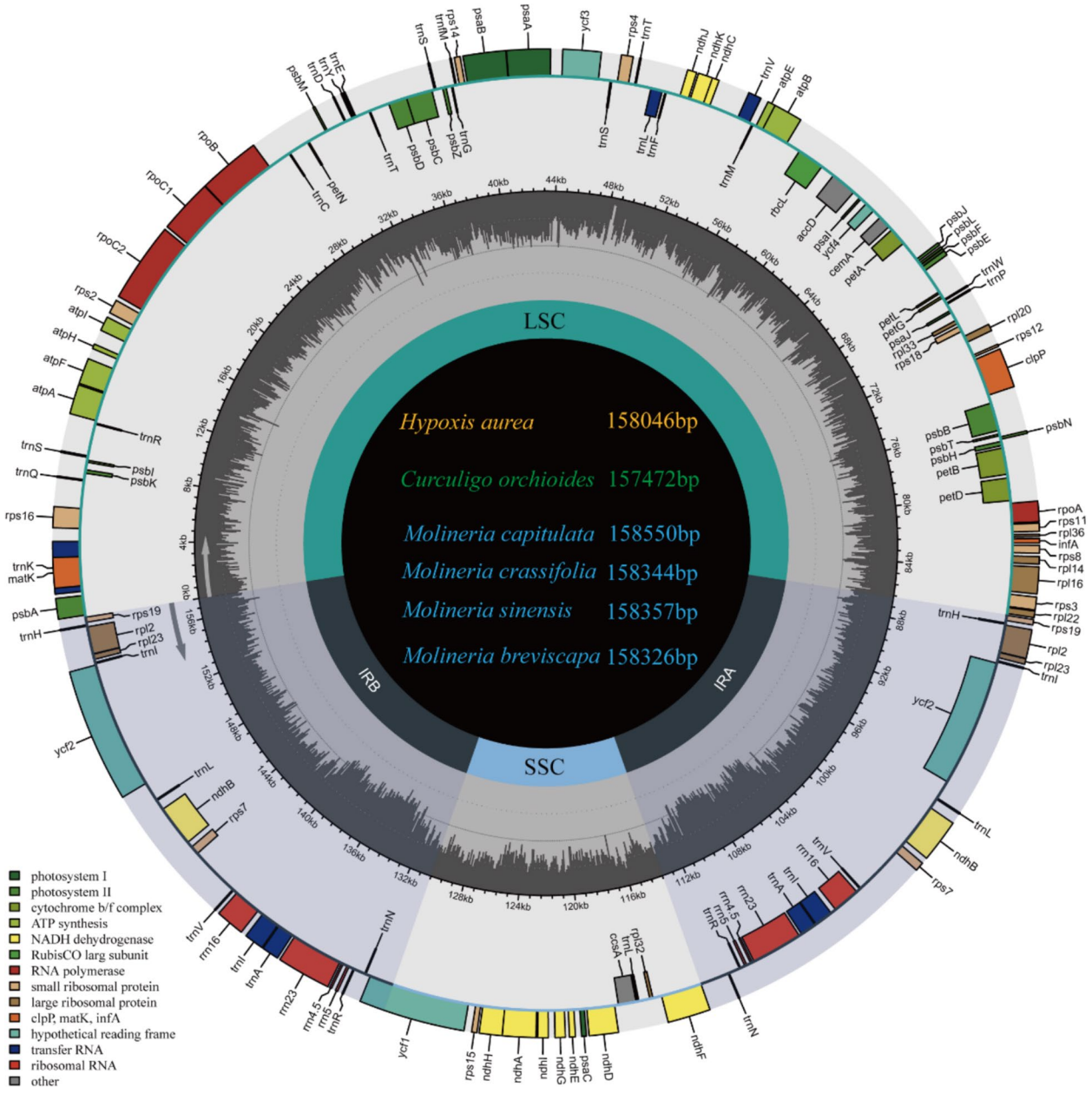

Fig. 1 Gene map of the chloroplast genome in six species from Hypoxidaceae. The whole genome length is 157,472–158,550 bp. Colored bars indicate different functional gene groups. The sector shading represents the inverted repeat regions (IRa and IRb), dividing cp genomes into small (SSC) and large (LSC) single–copy regions. The dark gray inner circle shows GC content

Four types of LSRs (long sequence repeats) were also identified based on six cp genomes from Hypoxidaceae, including complement (C), forward (F), palindromic (P), and reverse (R) (Fig. 2, Table S4). There were 48 (*Hypoxis aurea*) to 72 (*Curculigo orchioides*) LSR detected in the six cp genomes (Table S4). The length of LSRs was mainly in the range of 30 to 40 bp, featured by the smaller number of the longer LSRs. LSRs were primarily distributed in gene regions and least in intron regions (Fig. 2, Table S4).

Cluster analysis of the repeat sequence in the six cp genomes showed that they were clustered into three groups, including I (*Hypoxis aurea*), II (*Curculigo orchioides* – *Molineria capitulate*, and III (the other three *Molineria* species) (Fig. 2).

RSCU analysis

The relative synonymous codon usage (RSCU) of six cp genomes from Hypoxidaceae was calculated based on all protein–coding genes. It showed that six cp genomes contained 64 synonymous codons encoding 21 amino

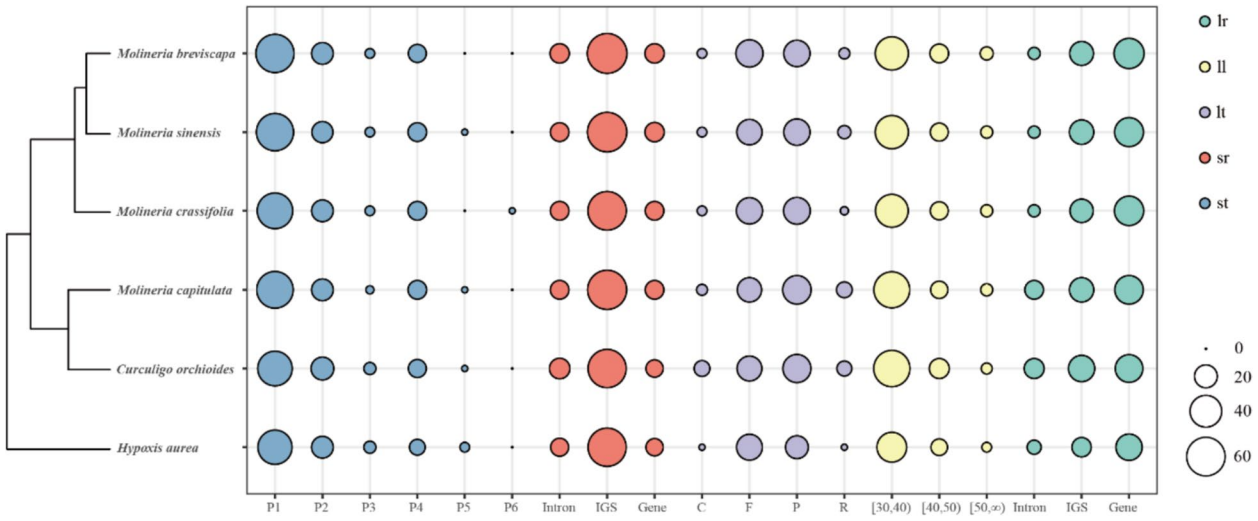

Fig. 2 Clustering tree and bubble map of chloroplast genome in six cp genomes from Hypoxidaceae based on Repeat sequence analysis. The color indicates the data type, and the size indicates the number. lr: regions with LSRs. ll: length of LSRs. lt: type of LSRs. sr: regions with SSRs. st: type of SSRs

Fig. 3 RSCU clustering tree and heat map of six chloroplast genomes from Hypoxidaceae. The color of the color block changes from purple to orange, indicating that the synonymous codon usage values are rising

acids (Fig. 3, Table S5). Among six cp genomes, leucine (Leu) possessed the highest number of codons (2,036–2,086), whereas tryptophan (Trp) had the lowest number of codons (367–375) (Table S5). According to the RSCU value, six cp genomes could be divided into two groups (Fig. 3), including 30 codons (RSCU < 1) and 34 codons (RSCU ≥ 1) found in *Hypoxis aurea* and *Curculigo orchioides*, and 32 codons (RSCU < 1) and 32 codons (RSCU ≥ 1) observed in four *Molineria* species. Almost all CDS had the standard ATG start codon in six cp genomes (Table S5), and TAA was the most common in three stop codons, with TGA and TAG also found (Table S5).

IR expansion and contraction

A comparative analysis of IR boundary positions and adjacent genes among the six cp genomes showed that they had a total of four boundaries (Fig. 4). In the six cp genomes, JLB (LSC/IRb) boundary was situated between *rpl22* gene and *rps19* gene, JSB (IRb/SSC) boundary between *trnN* gene and *ndhF* gene, while *ndhF* gene length has expanded into the IRb region by 2 to 11 bp. The JSA (SSC/IRa) boundary was found within *ycf1* gene, and its length expanded into the IRa region by 1306 to 1423 bp. The JLA (IRa/LSC) boundaries were located between *rps19* gene and *psbA* gene (Fig. 4). It showed that cp genome structure was consistent in four *Molineria* species, but distinct from those of other two genera

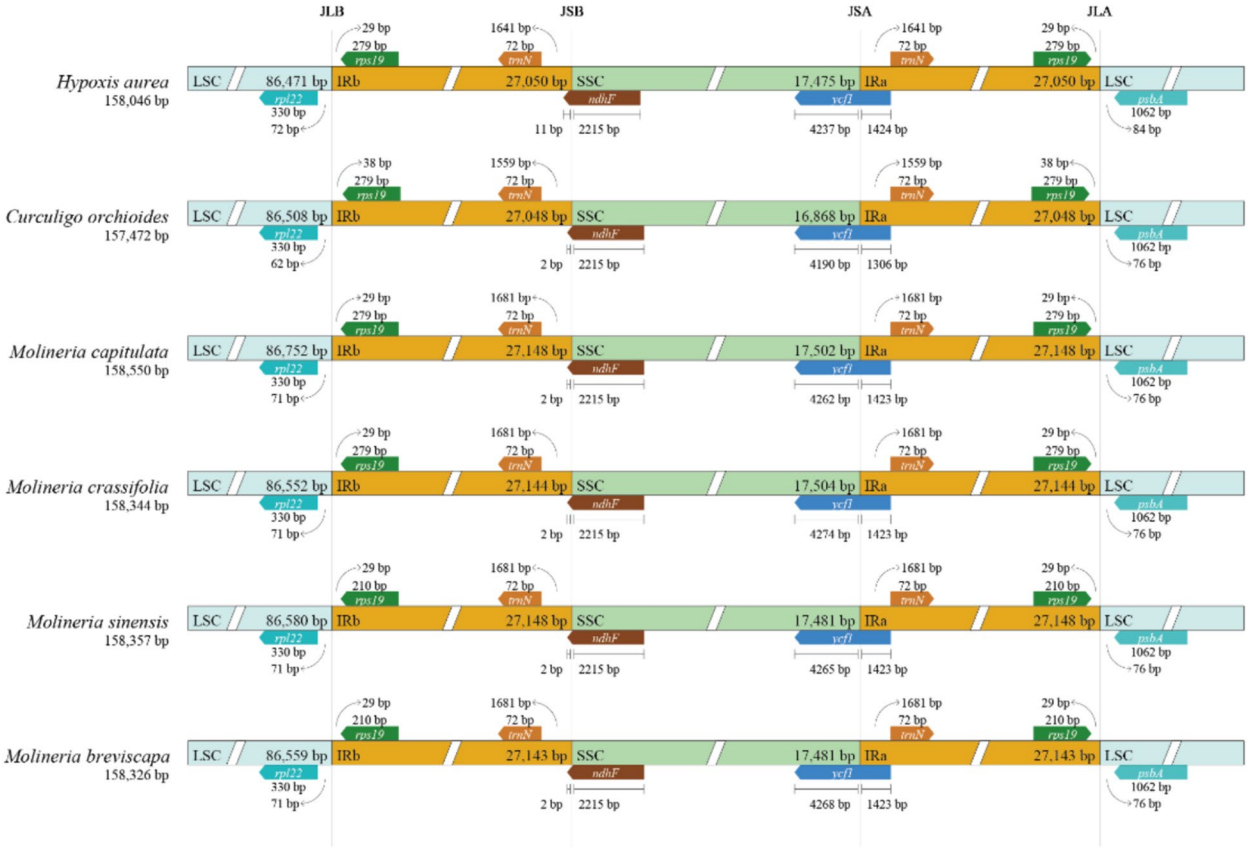

Fig. 4 A Comparative analysis of LSC, SSC, and IR region boundary among the six cp genomes from Hypoxidaceae. Genes adjacent to the junction were indicated as blocks of different colors

(*Hypoxis aurea* and *Curculigo orchioides*) by differences in the distance of the *rpl22*, *rps19*, and *psbA* genes from the border, and in the varied length of *ndhF* versus *ycf1* genes in the IR region.

Identification of divergence hotspot

To examine interspecific variations in six cp genomes from Hypoxidaceae by mVISTA Shuffle LANGAN model, *Hypoxis aurea* was designated as the reference sequence. It showed that IR regions were more stable than LSC and SSC regions which was related with the presence of the highly conserved rRNA genes (*rrn4.5*, *rrn5*, *rrn16*, *rrn23*) (Fig. 5). And then, a polymorphism analysis showed that Pi value in the coding regions ranged from 0 to 0.0202, with an average of 0.0061 (Fig. 6A, Table S6). There were three coding regions with Pi values exceeding 0.015, containing *petG*, *rpl33*, and *ycf1* (Fig. 6A, Table S6). And what's more, Pi value in the intergenic regions ranged from 0.001 to 0.0898, with an average of 0.0178. And there were two intergenic regions, where Pi value exceeded 0.04, namely, *trnC–GCA–petN* and *trnE–UUC–trnT–GGU* (Fig. 6B, Table S6).

Phylogenetic analysis

A maximum likelihood (ML) phylogenetic tree was reconstructed using complete chloroplast genome coding sequences (CDS) from 61 samples representing ten families of Asparagales (ingroup), including 16 samples spanning all 11 genera of Hypoxidaceae, with two taxa from Petrosaviales as the outgroup (Table S7). Due to the scarcity of complete chloroplast genomes across Hypoxidaceae lineages, we supplemented the analysis with CDS data from 10 additional taxa within each data-deficient genus (labeled as 'genus name + spp' in the tree; Table S7). It was indicated that members from each family were grouped into distinct, strongly supported clades (UFBoot: 96–100%), and then ten families were clustered into the 'higher asparagoid' and the 'lower asparagoid', respectively. The 'higher asparagoid', as the two core clades, was consisted of Asparagaceae *s.l.* and Amaryllidaceae *s.l.*. And the 'lower asparagoid', as the basal clades, was comprised by the remaining eight families, namely, Xanthorrhoeaceae Dumort. *s.l.*, Xeronemataceae M. W. Chase, Rudall & M. F. Fay, Iridaceae Juss., Doryanthaceae R. Dahlgren & Clifford + Tecophilaeaceae Leyb., Hypoxidaceae + Asteliaceae, and Orchidaceae (Fig. 7). The ML phylogeny strongly supported (UFBoot: 100%)

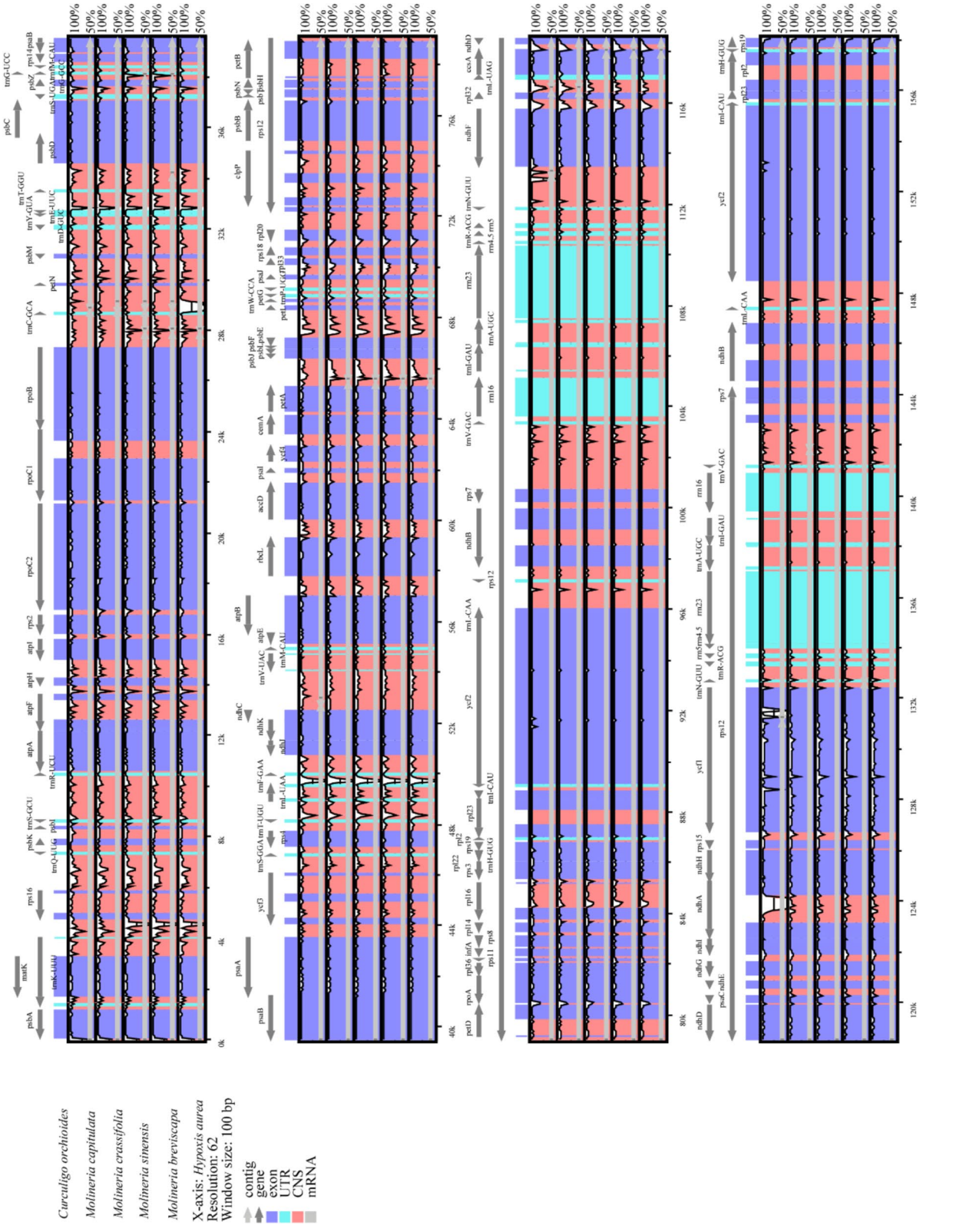

Fig. 5 Sequence alignment of six cp genomes from Hypoxidaceae by mVISTA Shuffle LANGAN model. Annotated genes are displayed along the top. The vertical scale represents the percent identity between 50 and 100%. Genome regions are color–coded as exon, intron, and conserved non–coding sequences (CNS)

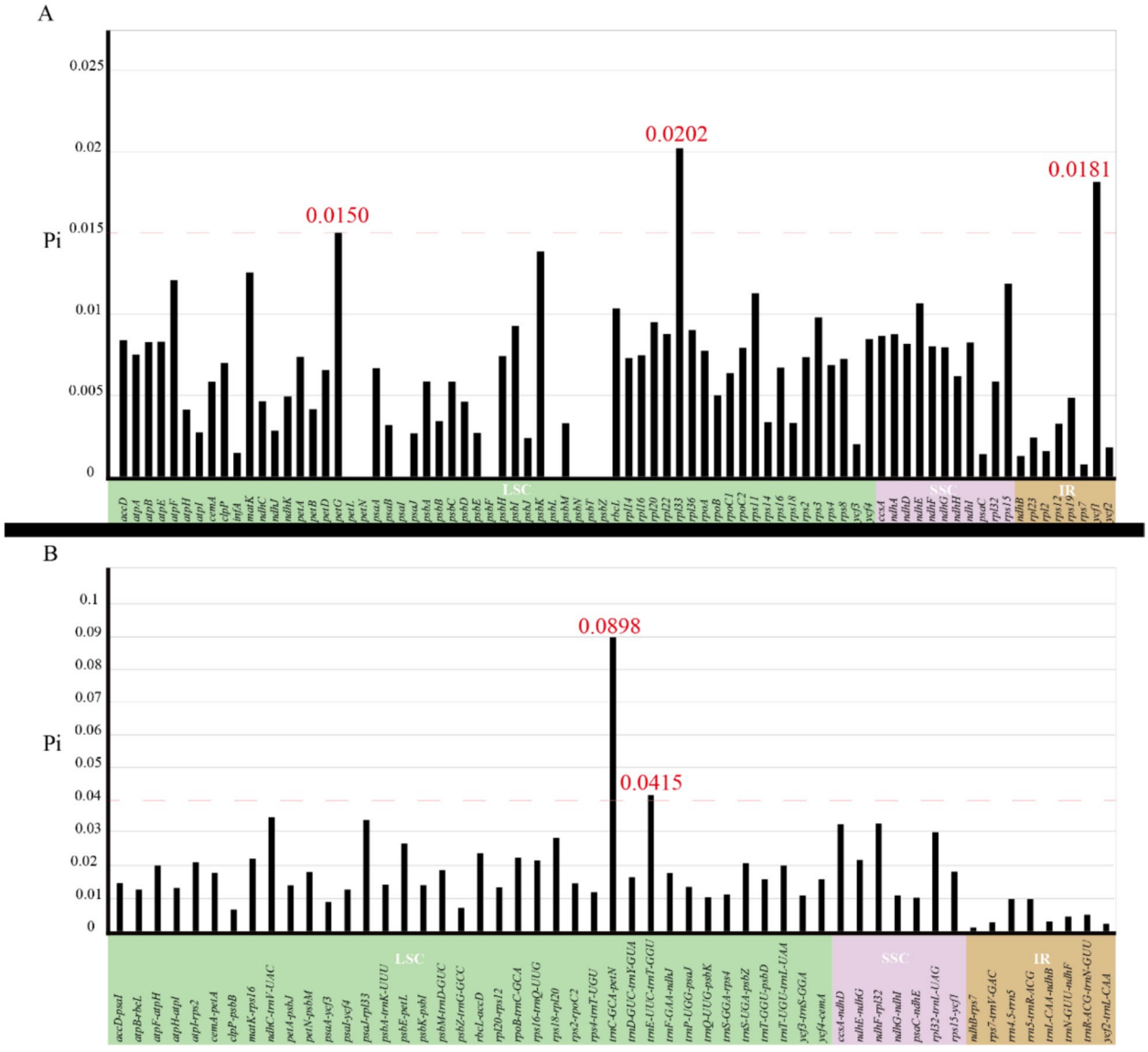

Fig. 6 Nucleotide polymorphism analysis of six cp genomes from Hypoxidaceae. **A** Comparison of the nucleotide polymorphism (Pi) among CDS regions. **B** Comparison of the nucleotide polymorphism among IGS regions. X–axis: position of the midpoint of a window; Y–axis: nucleotide polymorphism of each window. Highest variation hotspots for six chloroplast genomes were annotated on the graph. The colored blocks at the bottom delineate these gene locations in different regions

Orchidaceae as the earliest-diverging lineage and a sister to all other asparagoids. Within the remaining clades, Hypoxidaceae formed a sister relationship with Asteliaceae in the 'lower asparagoids,' while Amaryllidaceae s.l. and Asparagaceae s.l. were resolved as sister lineages in the 'higher asparagoids. Within Hypoxidaceae, all 16 sampled taxa formed a strongly supported monophyletic group (UFBoot: 100%), comprising three well-defined clades: a *Curculigo* clade (100% support), a *Hypoxis* clade (100% support), and a *Pauridia-Empodium* clade (78% support) (Fig. 7).

To further investigate the intergeneric relationships between *Curculigo* and *Molineria*, we reconstructed an additional ML phylogeny using four DNA regions (*rbc*L, *mat*K, *trn*L-*trn*F, and *trn*S-*trn*G) from 27 samples representing 26 species across five Hypoxidaceae genera (ingroup), with two Asteliaceae species as outgroup (Table S8). The analysis resolved these taxa into three strongly supported clades (Fig. 8): (1) a *Curculigo* clade (UFBoot: 100%), (2) a *Hypoxis* clade (UFBoot: 100%), and (3) a *Pauridia-Empodium* clade (UFBoot: 88%). The *Curculigo* clade comprised 21 species from five genera, the *Hypoxis* clade contained two Hypoxis species, and the *Pauridia-Empodium* clade included three species from *Pauridia* Harv. and *Empodium* Salisb. Within the *Curculigo* clade, three subclades were identified: a *Molineria*

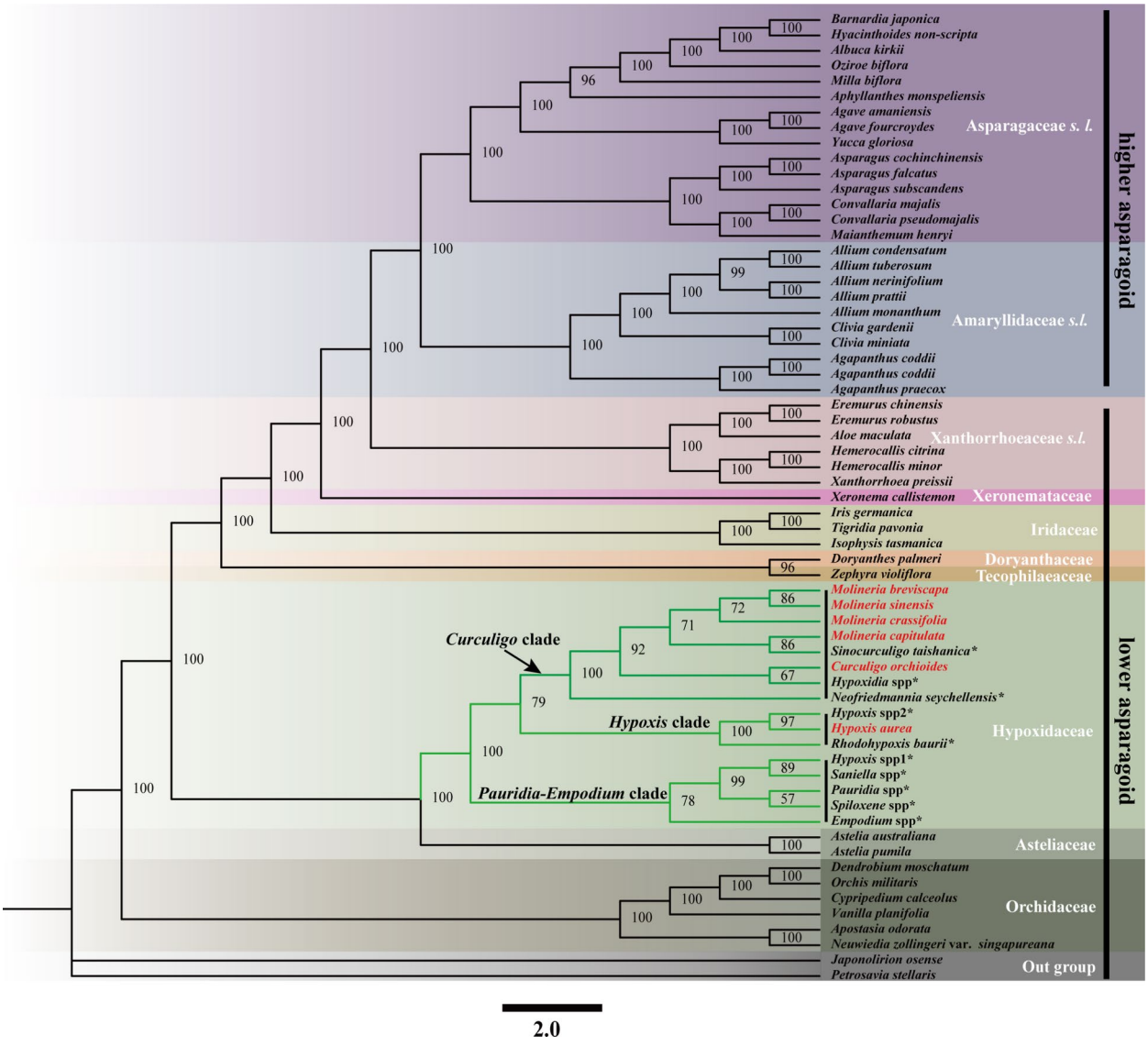

Fig. 7 A maximum likelihood phylogenetic tree reconstructed based on complete cp CDS sequences from 61 species representing ten families in Asparagales, with *Japonolirion osense* Nakai and *Petrosavia stellaris* Becc. as outgroup. *: only partial molecular fragments were used (see the 'Materials and methods' section). Numbers at nodes indicating ultrafast bootstrap values

subclade (UFBoot: 96%), a *Curculigo* subclade (UFBoot: 97%), and a Seychellean subclade (UFBoot: 87%). The Seychellean subclade was a sister to the other subclades (*Molineria* and *Curculigo*). The *Curculigo* subclade consisted of eight *Curculigo* species uniformly, and the Seychellean subclade included three Seychelles endemics (*Neofriedmannia seychellensis*, *Hypoxidia maheensis* F. Friedmann, and *H. rhizophylla* (Baker) F. Friedmann). However, the *Molineria* subclade contained seven *Molineria* species, two Borneo *Curculigo* species (*C. erecta* Lauterb. and *C. racemosa* Ridl.), and *Sinocurculigo taishanica*. It was further divided into two groups: Group I (UFBoot: 82%) included five *Molineria* species and *S. taishanica*, while Group II (UFBoot: 53%) comprised two East-Asian *Molineria* species and two Borneo *Curculigo* species.

Discussion

Systematic implication of cp genomic features in Hypoxidaceae

Chloroplast genomic data for Hypoxidaceae remain remarkably scarce, with only a single published genome (*Curculigo orchioides*) currently available [43]. Here, we conducted a comprehensive characterization and comparative analyses of six complete chloroplast genomes representing three taxonomically important genera within Hypoxidaceae from China: *Hypoxis* (*H. aurea*), *Curculigo* (*C. orchioides*), and *Molineria* (*M. capitulata*, *M. crassifolia*, *M. sinensis*, and *M. breviscapa*). This study

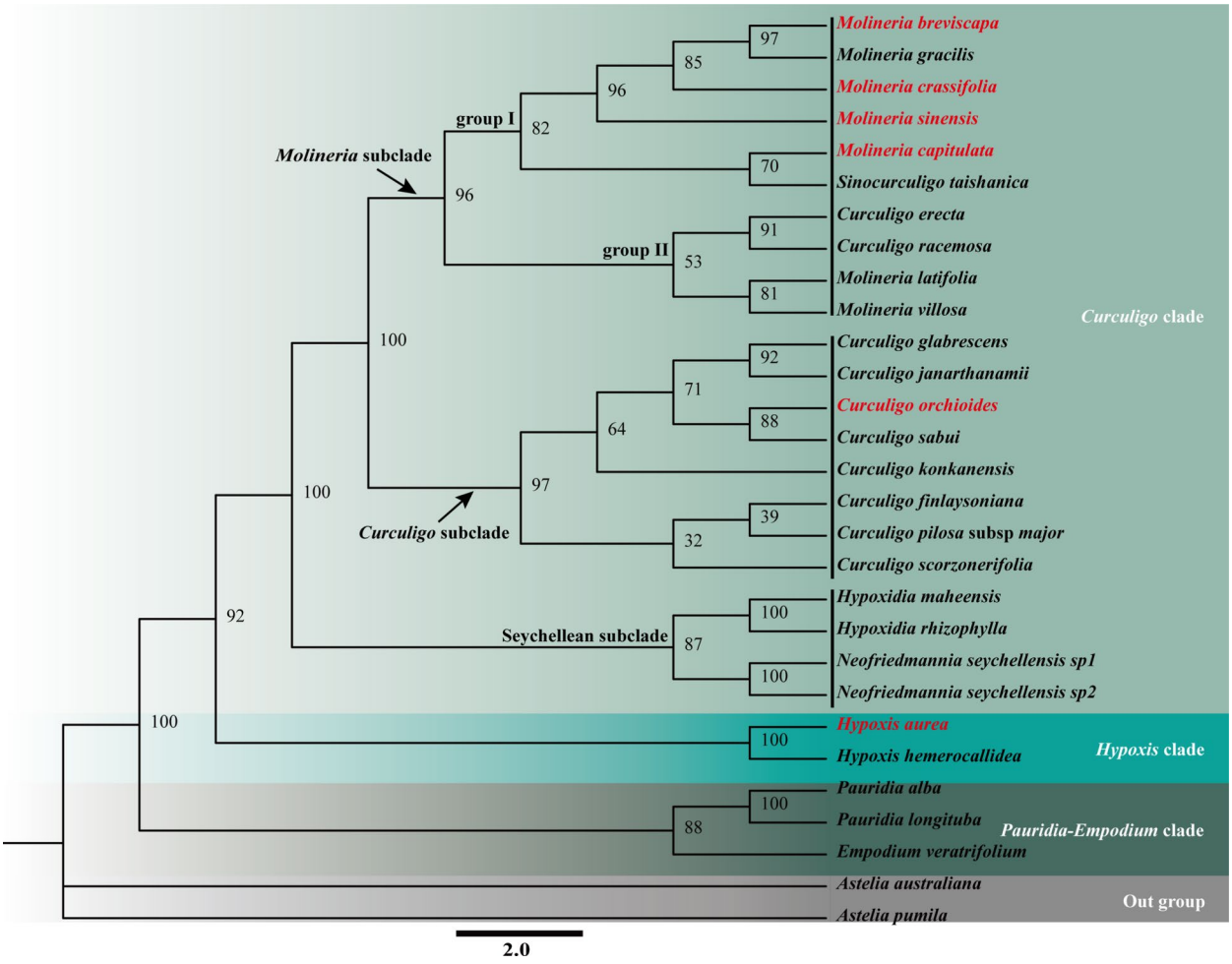

Fig. 8 A maximum likelihood phylogenetic tree reconstructed based on four DNA sequences (*rbcL*, *matK*, *trnL–trnF* and *trnS–trnG* intergenic regions) from 26 species from Hypoxidaceae, two species from *Astelia* Banks & Sol. ex R.Br. as outgroup. Numbers at nodes indicating ultrafast bootstrap values

represents the first systematic investigation of complete chloroplast genomes in Hypoxidaceae, providing new insights into both genome architecture and phylogenetic relationships within this understudied family.

The six cp genomes from Hypoxidaceae were highly conserved in the size, structure, and the number and content of gene. These six cp genomes displayed a typical quadripartite structure with the LSC and SSC regions partitioned by the IR regions, which were consistent with those of most angiosperms [35, 36, 39–42, 44–47]. The cp genomes were all AT–rich, and the GC contents in the IR region were significantly higher than that in other two regions, similar to those of other asparagoids [35, 36, 44–47]. This occurrence was possibly due to the four rRNA genes (rrn16, rrn23, rrn4.5, and rrn5) in the IR regions [48]. Gene content was highly conserved across the six cp genomes, each containing 132 genes (86 proteins–coding, 8 rRNA, and 38 tRNA genes), congruent with recent records in *Astelia* [44, 49].

Meanwhile, a complete functional set of *ndh* genes (*ndh A*/*B*/*C*/*D*/*E*/*F*/*G*/*H*/*I*/*J*/*K*) was observed in the six cp genomes from Hypoxidaceae, characteristic of *ndh*-complete genomes [50, 51]. While the *ndh* complex is known to play a crucial role in most photoautotrophic plants [50, 51], *ndh*-deleted genomes have been reported in certain Asparagales lineages, including *Strumaria truncata* Jacq. (Amaryllidaceae) [52] and some Orchidaceae species [35, 36]. Within Orchidaceae, the distribution of *ndh* genes shows remarkable variation at the subfamily level: Apostasioideae maintains *ndh*-complete genomes, Vanilloideae exhibits complete *ndh* deletions, while Cypripedioideae Lindl., Orchidoideae, and Epidendroideae display both *ndh*-complete and *ndh*-deleted patterns [35, 36, 53–56]. Phylogenetic evidence suggests that the common ancestor of orchids likely possessed a complete set of functional *ndh* genes [50]. The consistent presence of *ndh*-complete genomes in all six examined Hypoxidaceae species (representing three genera) provides important phylogenetic insights. This pattern distinguishes

Hypoxidaceae from both Amaryllidaceae (where it was previously classified [3, 4, 12, 13]) and Orchidaceae (with which it was once proposed to share close relationships [1, 14]).

Intergeneric discrimination of cp genomic features in Hypoxidaceae

The taxonomic distinction between *Curculigo* and *Molineria* has remained problematic, with unclear generic boundaries requiring further clarification [21, 24, 31]. To address this issue, we conducted comparative analyses of complete chloroplast genomes from six species representing three genera within Hypoxidaceae. Our study revealed several distinctive cp genomic features that may serve as potential molecular markers for intergeneric delimitation, including: (1) variations in repeat sequence composition and distribution, (2) patterns of codon usage bias, and (3) differential contractions and expansions in the inverted repeat (IR) regions.

Repeat sequence analyses (SSRs and LSRs) revealed their potential utility for generic delimitation among *Hypoxis*, *Curculigo*, and *Molineria* within Hypoxidaceae. Simple sequence repeats (SSRs), characterized by short tandem repeats (1–6 bp units), exhibited several advantageous features for genetic markers: high abundance, repetitive nature, structural simplicity, maternal inheritance through chloroplast genomes, and relative conservation [57–59]. Our analysis identified 82–92 SSRs across the six Hypoxidaceae chloroplast genomes. These SSRs were predominantly composed of poly-A or poly-T repeats, with mononucleotide repeats being the most abundant form—a pattern consistent with other asparagoid species [36, 46, 47]. Additionally, we detected 48–72 long sequence repeats (LSRs) that showed intergeneric variation. Cluster analysis based on repeat sequences clearly separated the six species into three distinct groups (Fig. 2): Group I: *Hypoxis aurea* (single species cluster); Group II: *Curculigo orchioides* and *Molineria capitulata*; Group III: The remaining three *Molineria* species.

Codon usage bias has proven to be a valuable molecular marker for systematic and evolutionary studies across diverse organisms [60, 61]. Our analysis of the six Hypoxidaceae chloroplast genomes revealed that: (1) all protein-coding sequences (CDS) contained the standard ATG start codon, consistent with the highly conserved pattern observed in other angiosperms [39–42]; and (2) the 64 codons encoding 21 amino acids showed distinct patterns of codon usage bias among the three genera. Notably, relative synonymous codon usage (RSCU) analysis revealed clear phylogenetic differentiation, separating the taxa into two distinct groups, namely, Group I (*Hypoxis aurea* and *Curculigo orchioides*), Group II (All four *Molineria* species (*M. capitulata*, *M. crassifolia*, *M. sinensis*, and *M. breviscapa*).

The inverted repeat (IR) regions represented the most conserved sections of the six Hypoxidaceae chloroplast genomes, consistent with patterns observed in most angiosperms [35, 36, 39, 40, 44, 45]. However, IR boundary regions exhibited significant expansions and contractions, which contribute to chloroplast genome evolution, serve as primary determinants of genome size variation [62, 63], and provide valuable genomic markers for understanding angiosperm taxonomic relationships [48, 64]. Our analysis revealed distinct patterns of IR boundary variation among the studied taxa. The four *Molineria* species (*M. capitulata*, *M. crassifolia*, *M. sinensis*, and *M. breviscapa*) shared identical IR boundaries. Both *Hypoxis aurea* and *Curculigo orchioides* displayed unique IR boundary configurations that differed from each other and from the *Molineria* species.

Five DNA barcodes for authentication of medicinal hypoxid species

The rhizomes of Hypoxidaceae species, including *Curculigo orchioides* (known as 'Xianmao' in traditional Chinese medicine) [6], *Molineria capitulata*, *M. breviscapa*, and *Hypoxis aurea* [7–9], are widely used in traditional medicine throughout China and Southeast Asia. However, inadequate taxonomic characterization has led to frequent misidentification and misuse of these medicinal resources in commercial markets [9–11]. Our nucleotide polymorphism analysis identified three protein-coding regions (*pet*G, *rpl*33, and *ycf*1) with exceptionally high Pi values across the six chloroplast genomes. The ycf1 gene, like the established barcode marker *mat*K, has demonstrated particular utility for both medicinal plant authentication [65–67] and phylogenetic reconstruction [2, 21]. Additionally, two intergenic spacer regions (*trnC–GCA–petN* & *trnE–UUC–trnT–GGU*) exhibited substantial sequence variability. These findings suggest five promising molecular markers for authenticating Hypoxidaceae medicinal materials, containing Protein-coding genes (*petG*, *rpl33*, and *ycf1*) and Intergenic spacers (*trnC–GCA–petN* & *trnE–UUC–trnT–GGU*).

Phylogenetic significance of complete cp genome in Hypoxidaceae

The phylogeny of asparagoid families based on complete chloroplast genomes was presented here, incorporating 61 samples representing 42 genera across all ten recognized families. These samples were six newly sequenced Hypoxidaceae chloroplast genomes; 45 publicly available genomes from nine other asparagoid families. The resulting topology showed strong congruence with previous molecular studies [22, 23], but with significantly improved nodal support (UFBoot: 96–100%) compared to analyses using limited molecular markers. The monophyly of astelioid clade (UFBoot: 100%) was confirmed

with a robust support, while received relatively low support values by limited molecular sequence data analyses [22, 24], and by the 77 plastid gene analyses of Givnish et al. [23]. These results demonstrate the superior phylogenetic resolution provided by complete chloroplast genomes compared to traditional molecular markers [22–24, 68].

The systematic position of Hypoxidaceae has long been debated, with conflicting classifications placing its members either as a tribe within Amaryllidaceae [3, 4, 13] or as a potential sister group to Orchidaceae based on morphological similarities [15, 16, 18]. Our comprehensive phylogenetic analysis of chloroplast coding sequences provides conclusive evidence that Hypoxidaceae is a strongly supported monophyletic clade (UFBoot: 98%), being a sister with Asteliaceae within the 'lower asparagoid' lineage. These findings definitively confirmed Hypoxidaceae as phylogenetically distinct from both Amaryllidaceae s.l. (which clusters with Asparagaceae s.l.) and Orchidaceae.

Meanwhile, in order to find some clues for the phylogenetic implications of cp genome from Hypoxidaceae, an ML tree (Fig. 7) was reconstructed based on the six complete cp genomes from China and the concatenated chloroplast gene sequences (2–70 genes) as complete chloroplast genomes from ten taxa representing eight genera, with an effort to encompassing main distribution of the Southern Hemisphere. The phylogenetic topology showed strong congruence with previous molecular studies [22, 23], but with significantly improved nodal support (UFBoot: 96–100%) compared to analyses using limited molecular markers. Three well-supported subclades were identified, including a *Curculigo* clade (UFBoot: 100%), a *Hypoxis* clade (UFBoot: 100%), a *Pauridia-Empodium* clade (UFBoot: 78%). The *Curculigo-Molineri*a complex showed topological congruence with previous molecular studies [21, 24], and the phylogenetic position of *Neofriedmannia seychellensis* remains ambiguous requiring further research. However, a clear generic boundary between *Curculigo* and *Molineria* could be detected (Fig. 8), despite limited taxon sampling (four *Molineria* spp., one *Curculigo* sp., and an uncertain position of *Sinocurculigo taishanica*) [21–24]. And then, this generic delimitation between these two genera is to be discussed as below.

Phylogenetic analysis of Curculigo clade

The intergeneric phylogeny among Hypoxidacae has been unclear and needed more data for clarification, with a focus on the delimitation between *Curucligo* and *Monlineria* [2, 19, 21, 24, 25, 28–30]. Our comprehensive phylogenetic analysis of Hypoxidaceae, incorporating four chloroplast DNA markers and expanded taxon sampling, provides significant insights into the intergeneric relationships within this family, particularly concerning the taxonomically challenging *Curculigo* clade. It incorporated sequence data from five additional species (*M. breviscapa*, *C. glabrescens*, *C. janarthanamii* Gore & S. P. Gaikwad, *C. sabui* S. P. Gaikwad & Gore, and *C. konkanensis* Chandore, Mane & Borude) to augment previous datasets [21, 24]. The robust phylogenetic framework presented here (Fig. 8) confirms and extends previous findings [2, 21, 22, 30, 31] while revealing new complexities in generic delimitation between *Curculigo* and *Molineria* [2, 21, 24]. Three major clades identified in our study—*Curculigo*, *Hypoxis*, and *Pauridia-Empodium*—show complete concordance with both our complete chloroplast genome data and prior molecular studies [2, 21, 22, 30, 31]. This consistency across different datasets and analytical approaches strongly supports the fundamental division of Hypoxidaceae into these three lineages [21, 22, 30, 31].

In the *Curculigo* clade, three distinct subclades with strong supports were identified, namely, *Molineria* subclade (UFBoot: 96%), *Curculigo* subclade (UFBoot: 97%), and Seychellean subclade (UFBoot: 87%) (Fig. 8). The Seychellean subclade contained three Seychelles endemic species (*Neofriedmannia seychellensis, Hypoxidia maheensis*, and *H. rhizophylla*). This subclade was supported as a monophyly and adopted by previous findings, although *Neofriedmannia seychellensis* was removed from *Curculigo* and its phylogenetic position remains ambiguous [21, 24]. The *Curculigo* subclade was composed eight *Curculigo* species consistently, while the *Molineria* subclade consisted of seven *Molineria* species, two Borneo *Curculigo* species (*C. erecta* Lauterb. and *C. racemosa* Ridl.), and a Chinese endemic *Sinocurculigo taishanica.* The presence of two Borneo *Curculigo* species (*C. erecta* and *C. racemosa*) within the *Molineria* Subclade, along with the ambiguous placement of *Sinocurculigo taishanica*, highlights persistent challenges in generic delimitation.

Our phylogenetic analysis reveals that the taxonomic ambiguity primarily stems from the inclusion of two Borneo *Curculigo* species within the *Molineria* subclade, the phylogenetic position of *Sinocurculigo taishanica*, and the moderate supports within the *Molineria* subclade (Fig. 8, Groups I & II). Group I was consisted of five *Molineria* species and a new species of *Sinocurculigo taishanica* with a strong support (UFBoot: 82%). It was noteworthy that *Sinocurculigo taishanica* was thought to be treated as a separated genus, or as a member of *Molineria,* when it was firstly described and to be close to *Molineria capitulata* [2]. According to our phylogenetic analysis, *Sinocurculigo taishanica* could be better placed in *Molineria* than to be in a monotypic genus. Meanwhile, two other *Molineria* species and two Borneo *Curculigo* species were clustered into the group II with

a low support (UFBoot: 53%) which was consistent with the other studies [21, 24]. And the data on these four taxa were scarce and unavailable for further research. These results revive the century–old debate regarding: merging *Curculigo* and *Molineria* [21], or alternatively, namely, transferring *Sinocurculigo taishanica* and the two Borneo *Curculigo* species into *Molineria* Subclade. The latter suggestion follows the same logic applied to the recognition of the Seychellean subclade as distinct despite its uncertain position, maintaining parallel standards of generic delimitation across comparable clades [21, 24]. What's important is that the cp genome features are conserved in the four *Molineria* speicies, distinct from *Curculigo* and *Hypoxis*, indicating a generic variation systematically. Therefore, the phylogenetic evidence presented herein, combined with distinct chloroplast genomic features, support *Molineria* subclade being separated from *Curculigo* subclade. It provides an alternatively robust resolution for clarifying the generic delimitation between *Curculigo* and *Molinera* by reassigning certain contentious species to *Molineria* Subclade — thereby maintaining its monophyly.

Our findings also reveal significant issues with current species circumscriptions, particularly regarding *Curculigo glabrescens* and *Molineria villosa*. Both of them were considered as synonyms of the same species of *M. latifolia* Dryand. ex W.T. Aiton without any reason in some records [4, 69, 70]. However, species of the same genus (*C. glabrescens*, *M. villosa* and *M. latifolia*), respectively, clustered into *Curculigo* and *Molineria* subclades (Fig. 8). The phylogenetic separation of these taxa, despite previous treatments as conspecific with *M. latifolia* [4, 69, 70], providing an urgent need for comprehensive taxonomic revision incorporating both molecular and morphological data.

Conclusion

The complete cp genomes six species of three genera from Hypoxidaceae were sequenced and analyzed in detail, including the general data on the genome, repeat sequence, codon usage, IR expansion and contraction, structural comparison and divergence hotspot identification analyses, and phylogenetic analysis. A comparative analysis showed that cp genomes were highly consistent of four *Molineria* species, while varied at the generic level, such as *Hypoxis aurea* and *Curculigo orchioides*. Five molecular markers (*psbK–psbI*, *rpoB–trnC, ndhF–rpl32, ycf1,* and *trnE–trnT*) were selected for authentication of Hypoxidaceae medicinal materials. Hypoxidaceae is a monophyletic clade, containing three major clades, being a sister to Asteliaceae based on cp genomic data. It showed that current generic boundaries in Hypoxidaceae, particularly between *Curculigo* and *Molineria*, require reevaluation. The phylogenetic evidence presented here, combined with distinct chloroplast genome features, supports the recognition of *Molineria* subclade distinct from *Curculigo* subclade and maintaining the monophyly by transferring select species into *Molineria*. Further genomic and morphological data, incorporating extensive sampling across the Southern hemisphere, would significantly enhance our understanding of phylogenetic relationships within Hypoxidaceae.

Materials and methods

Ethical statement

No specific permits were required for the collection of specimens for study, and all procedures were conducted in compliance with relevant Chinese laws.

Plant materials and chloroplast genome sequencing

Plants of six species investigated were cultivated in Botanical Garden of Xishuangbanna South Medicine, Chinese Academy of Medical Science, Jinghong, China (Fig. 9). Prof. Lu Li (lilu@swfu.edu.cn) identified these six species sampled by their habitat, leaf, flower, and fruit morphological features. These specimens and samples were collected and observed by Wang Yunqiang, Tian Qin, and Duan Hanning. Voucher specimens for six species (*Hypoxis aurea*, *Curculigo orchioides*, *Molineria capitulata*, *M. crassifolia*, *M. sinensis*, and *M. breviscapa*) have been deposited at the Herbarium of Southwest Forestry University (HSFU) under accession numbers Lilu2020001 to Lilu2020006 respectively. Genomic DNA of each sample was extracted from the silica gel–dried leaf tissues using the modified CTAB method with the TiangenDNA kit (TIANGEN, China) [48]. Paired–end libraries with an average insert size of approximately 400 bp were prepared using a TruSeq DNA Sample Prep Kit (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's instructions. The libraries were sequenced on the Illumina HiSeq 2500 platform at Personalbio (two times 150 bp; Illumina, Shanghai, China). Raw data were filtered using Fastp v0.23.1 (https://github.com/OpenGene/fastp) to obtain high–quality reads by the sliding window method to drop the low–quality bases of each read's head and tail [71].

Chloroplast genome assembly and annotation

The six complete cp genomes from the clean reads were assembled by the GetOrganelle version 1.7.7.1 (https://github.com/Kinggerm/GetOrganelle) [72]. The assembled cp genomes were annotated using the online program CPGAVAS2 (http://47.96.249.172:16019/analyzer/home) [73], and further manually adjusted and confirmed using Geneious Prime version 2024.0.5 (https://www.geneious.com/features/prime). The complete cp genomes sequences of six species were submitted to GenBank (Accession numbers: NC–052741, NC–053892 and

Fig. 9 Plant morphology of six species from Hypoxidaceae. **A** *Hypoxis aurea*. **B**, **C** *Curculigo orchioides*. **D**, **E** *Molineria capitulate*. **F**–**H** *M. crassifolia*. **I**, **J** *M. sinensis*. **K**, **L** *M. breviscapa*

PV738957 to PV738960). The circular genome maps were drawn by the OGDRAW program (https://chlorobox.mpimp-golm.mpg.de/OGDraw) [74].

Comparative analysis

By setting the minimum number of repeats to 10, 5, 4, 3, 3, and 3 for mononucleotide (P1), dinucleotide (P2), trinucleotide (P3), tetranucleotide (P4), pentanucleotide (P5), and hexanucleotide (P6), respectively, simple sequence repeats (SSR), a tract of repetitive DNA that typically ranges in length from 1 to 6 nucleotides, were detected via MISA (https://webblast.ipk-gatersleben.de/misa/index.php?action=1) [75]. The repetitive structures, repeat sizes, and locations of forward match (F), reverse match (R), palindromic match (P), and complementary match (C) nucleotide repeat sequences were identified by REPuter v2.74 (https://bibiserv.cebitec.uni-bielefeld.de/reputer/) [76], with minimal repeat size set to 30 bp, hamming distance set to 3 and maximum computed repeats set to 90.

Condon usage was analyzed by MEGA11 software (https://www.megasoftware.net/dload_win_beta) [77], and the relative synonymous codon usage (RSCU) and amino acid frequencies were calculated with default settings [78]. RSCU clustering tree and heat map were drawn using the R package (pheatmap). The IR region analysis was performed using JSHYCloud (http://cloud.genepioneer.com:9929). The pairwise alignments and sequence divergence using the mVISTA with Shuffle–LAGAN mode (https://genome.lbl.gov/cgi–bin/VistaInput?nu

m_seqs=2) [79]. Highly variable sites were identified by DnaSP V6 software (https://gitcode.com/open-source-toolkit/658b6/?utm_source=tools_gitcode&index=top&type=card) [80].

Phylogenetic analysis

For phylogenetic analysis, the complete chloroplast CDS sequences of 61 representative species from Asparagales were selected (Table S7). Two species from Petrosaviales were selected as outgroups because it's the sister clade of Asparagales [70]. Due to the scarcity of complete chloroplast genome data for most Hypoxidaceae taxa, we utilized chloroplast coding sequence (CDS) data from various species within genus, which are labeled in the phylogenetic tree as 'genus name + spp' (Table S7). These CDS sequences were extracted by PhyloSuite version 1.2.3 (https://github.com/dongzhang0725/PhyloSuite) [81, 82], aligned by MAFFT version 7 [83], trimmed by Gblocks, and concatenated by plugins in PhyloSuite version 1.2.3 [81, 82]. The Maximum-Likelihood (ML) tree was constructed using IQ-TREE 2 (https://github.com/iqtree/iqtree2) with the GTR + F + I + G4 model, which was identified as the best-fit substitution model through ModelFinder (implemented in IQ-TREE) based on Bayesian Information Criterion (BIC) scores, with 5000 ultrafast bootstrap (UFBoot) [84–86].

In order to further explore the phylogenetic relationships of Hypoxidaceae, a phylogenetic tree of Hypoxidaceae was constructed based on *rbcL*, *matK*, *trnL–trnF* and *trnS–trnG* intergenic regions of 26 species from Hypoxidaceae as ingroup (including all available data from the *Curculigo* clade), with two species from Asteliaceae as outgroup (because it's the sister clade of Hypoxidaceae) using ML method (Table S8). The four DNA sequences, were trimmed by Gblocks, and concatenated by plugins in PhyloSuite version 1.2.3 [81, 82]. The phylogenetic tree was constructed by IQ–TREE 2 in K3Pu + F + G4 model, which was identified as the best-fit substitution model through ModelFinder (implemented in IQ-TREE) based on Bayesian Information Criterion (BIC) scores, with 5000 ultrafast bootstrap (UFBoot) [84–86].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-07191-5.

Supplementary Material 1.

Acknowledgements

We thank Dr. Kaifeng Tao for suggestions on data analysis, and Associate Professor Yuxiao Zhang for providing the computer server.

Authors' contributions

L.L. and Y.L. designed and supervised the research. Q.T., Y.W., and H.D. collected all plant samples. L.L. identified all plant samples. D.M., Q.T., and Y.Z. performed the experiments, assembled and annotated the genomes, and analyzed the data. D.M. drafted the manuscript. L.L. and Y.L. revised the manuscript. All authors have read and approved the final manuscript.

Funding

This research was funded by National Nature Science Foundation of China (NSFC 32060049 and NSFC 32270225).

Data availability

The raw data generated in this research are available in the NCBI BioProject (PRJNA667995 and PRJNA1273022).

Declarations

Ethics approval and consent to participate

The plant materials used in this research comply with relevant institutional, national, and international guidelines and legislation.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Clinical trial number

Not applicable.

Author details

[1]College of Forestry, Southwest Forestry University, Kunming, Yunnan 650224, China
[2]Yunnan Academy of Biodiversity, Southwest Forestry University, Kunming, Yunnan 650224, China
[3]Honghe Tropical Agriculture Institute of Yunnan, Hekou, Yunnan 661300, China
[4]Key Laborartory of Dai and Southern Medicine of Xishuangbanna Dai Autonomous Prefecture, Jinghong 666100, China
[5]CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, China
[6]Southeast Asian Biodiversity Research Institute, Chinese Academy of Sciences & Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan 666303, China

Received: 8 June 2025 / Accepted: 6 August 2025
Published online: 30 September 2025

References

1. Kubitzki K, Huber H, editors. Flowering plants, Monocotyledons: Lilianae (except Orchidaceae). Berlin; New York: Springer-Verlag; 1998.
2. Liu KW, Xie GC, Chen LJ, Xiao XJ, Zheng YY, Cai J, et al. *Sinocurculigo*, a new genus of Hypoxidaceae from China based on molecular and morphological evidence. PLoS One. 2012;7:e38880.
3. Tsi Z, W. Meerow A. *Hypoxis*. In: Flora of China. Science Press; Beijing. 2000. p. 273.
4. Tsi Z, W. Meerow A. *Curculigo*. In: Flora of China. Science Press; Beijing. 2000. p. 273.
5. Remizowa MV, Sokoloff DD, Rudall PJ. Evolutionary history of the monocot flower. Ann Mo Bot Gard. 2010;97:617–45.
6. National Pharmacopoeia Committee. Pharmacopoeia of the People's Republic of China. Beijing: China Medical Science and Technology Press; 2020.
7. Editorial Board of China Herbal. Chinese herbal medicine. In: China herbal. Shanghai: Shanghai Science and Technology Press; 1999.
8. Wang Y, Li J, Li N. Phytochemistry and pharmacological activity of plants of genus *Curculigo*: an updated review since 2013. Molecules. 2021;26:3396.
9. Nie Y, Dong X, He Y, Yuan T, Han T, Rahman K, et al. Medicinal plants of genus *Curculigo*: traditional uses and a phytochemical and ethnopharmacological review. J Ethnopharmacol. 2013;147:547–63.
10. Ncube B, Ndhlala AR, Okem A, Van Staden J. *Hypoxis* (Hypoxidaceae) in African traditional medicine. J Ethnopharmacol. 2013;150:818–27.

11. Chen QS, Chen WQ, Yang SY. Pharmacologic study of *Curculigo orchioides* Gaertn. China J Chin Mater Med. 1989;14:618–20.
12. Pax F. Amaryllidaceae., Die naturlichen Pflanzenfamilien. Leipzig. 1930;15:391–431.
13. Abd El-Ghany S, Naga G, Sharaawy S, Abouseadaa H. Taxonomic relationships of some taxa of the family Amaryllidaceae (s.l.) as reflected by macro-and micromorphological characters. Taeckholmia. 2023;43:141–76.
14. Monocotyledons John H. The families of flowering plants. London: Macmillan; 1934.
15. Garay LA. On the origin of the Orchidaceae. Bot Mus Leafl Harv Univ. 1960;19:57–96.
16. Garay LA. On the origin of the Orchidaceae, part II. J Arnold Arbor. 1972;53:202–15.
17. K. Zimmerman J, Ackerman JD. Orchid taxonomy, ecology, and evolution. In: The first 100 years of research on Barro Colorado: plant and ecosystem science. Smithsonian Institution Scholarly Press;Washington. 2024.
18. Perez-Escobar OA, Bogarín D, Przelomska NAS, Ackerman JD, Balbuena JA, Bellot S, et al. The origin and speciation of orchids. 2023.
19. Rudall PJ, Chase MW, Cutler DF, Rusby J, Bruijn AYD. Anatomical and molecular systematics of Asteliaceae and Hypoxidaceae. Bot J Linn Soc. 1998;127:1–42.
20. Takhtajan AL. Flowering plants: origin and dispersal. Edinburgh: Oliver & Boyd; 1969.
21. Kocyan A, Snijman DA, Forest F, Devey DS, Freudenstein JV, Wiland-Szymańska J, et al. Molecular phylogenetics of Hypoxidaceae – evidence from plastid DNA data and inferences on morphology and biogeography. Mol Phylogenet Evol. 2011;60:122–36.
22. Chen S, Kim DK, Chase MW, Kim JH. Networks in a large-scale phylogenetic analysis: reconstructing evolutionary history of Asparagales (Lilianae) based on four plastid genes. PLoS One. 2013;8:e59472.
23. Givnish TJ, Zuluaga A, Spalink D, Soto Gomez M, Lam VKY, Saarela JM, et al. Monocot plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of monocots. Am J Bot. 2018;105:1888–910.
24. Birch JL, Kocyan A. Biogeography of the monocotyledon astelioid clade (Asparagales): a history of long-distance dispersal and diversification with emerging habitats. Mol Phylogenet Evol. 2021;163:107203.
25. Rudall PJ. Unique floral structures and iterative evolutionary themes in Asparagales: insights from a morphological cladistic analysis. Bot Rev. 2002;68:488–509.
26. Zuntini AR, Carruthers T, Maurin O, Bailey PC, Leempoel K, Brewer GE, et al. Phylogenomics and the rise of the angiosperms. Nature. 2024;629:843–50.
27. Pires C, Maureira I, Givnish T, Systma K, Seberg O, Peterson G, et al. Phylogeny, genome size, and chromosome evolution of Asparagales. Aliso. 2006;22:287–304.
28. Hilliard OM, Burtt BL. Notes on some plants from southern Africa chiefly from Natal: VII. 1978.
29. Kocyan A, Endress PK. Floral structure and development and systematic aspects of some "lower" Asparagales. Plant Syst Evol. 2001;229:187–216.
30. Rudall PJ, Furness CA, Chase MW, Fay MF. Microsporogenesis and pollen sulcus type in Asparagales (Lilianae). Can J Bot. 1997;75:408–30.
31. Kocyan A, Wiland-Szymańska J. *Friedmannia*: a new genus from the Seychelles and the beginning of a generic realignment of *Curculigo* (Hypoxidaceae). Phytotaxa. 2016;283:54.
32. Kocyan A, Wiland-Szymańska J. A new name and a new combination for *Friedmannia* nom. illeg. (Hypoxidaceae). Phytotaxa. 2017;291:239.
33. Wang S, Wang Y, Chen Y, Zhao Y, Gao M. Advances in the evolutionary mechanisms and genomic studies of sexual differentiation in Lauraceae plants. Int J Mol Sci. 2025;26:4335.
34. Zhang W, Wang H, Dong J, Zhang T, Xiao H. Comparative chloroplast genomes and phylogenetic analysis of *Aquilegia*. Appl Plant Sci. 2021;9:e11412.
35. Tao K, Tao L, Huang J, Duan H, Luo Y, Li L. Complete chloroplast genome structural characterization of two *Aerides* (Orchidaceae) species with a focus on phylogenetic position of *Aerides flabellata*. BMC Genomics. 2024;25:552.
36. Tao K, Tang L, Luo Y, Li L. Complete chloroplast genome of eight *Phaius* (Orchidaceae) species from China: comparative analysis and phylogenetic relationship. BMC Plant Biol. 2025;25:37.
37. Tamboli AS, Shelke PE, Chougule RN, Youn JS, Choo YS, Pak JH, et al. Comparative chloroplast genomics and phylogeny of *Crinum brachynema* (Amaryllidaceae): a narrow endemic geophyte of the western Ghats of India. Plant Mol Biol Rep. 2025;43:1283–96.
38. 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;112:1454–1463.
39. Nguyen VB, Linh Giang VN, Waminal NE, Park HS, Kim NH, Jang W, et al. Comprehensive comparative analysis of chloroplast genomes from seven *Panax* species and development of an authentication system based on species-unique single nucleotide polymorphism markers. J Ginseng Res. 2020;44:135–44.
40. Feng S, Zheng K, Jiao K, Cai Y, Chen C, Mao Y, et al. Complete chloroplast genomes of four *Physalis* species (Solanaceae): lights into genome structure, comparative analysis, and phylogenetic relationships. BMC Plant Biol. 2020;20:242.
41. Wu L, Nie L, Xu Z, Li P, Wang Y, He C, et al. Comparative and phylogenetic analysis of the complete chloroplast genomes of three *Paeonia* section *Moutan* species (Paeoniaceae). Front Genet. 2020;11:980.
42. Gao C, Wu C, Zhang Q, Zhao X, Wu M, Chen R, et al. Characterization of chloroplast genomes from two *Salvia* medicinal plants and gene transfer among their mitochondrial and chloroplast genomes. Front Genet. 2020;11:574962.
43. Deng G, Zhang R, Yang J, Deng C, Zhang Y. Characterization of the complete chloroplast genome of medical plant *Curculigo orchioides* Gaertn. (Amaryllidaceae). Mitochondr DNA B Resources. 2021;6:285–7.
44. Amor MD, Holmes GD, James EA. Characterization of the complete plastid genome of *Astelia australiana* (J. H. Willis) L. B. Moore (Asteliaceae, Asparagales). Mitochondr DNA B DNA Mapping Seq Anal. 2020;5:656–7.
45. González-Trillo AC, Reyes López MÁ, Almaraz-Abarca N, Herrera-Arrieta Y, Gutiérrez-Velázquez MV, Barraza Salas M, et al. Characterization of the complete chloroplast genome sequence of *Agave durangensis* (Asparagales: Asparagaceae: Agavoideae). Mitochondr DNA B. 2024;9:536–40.
46. Kamra K, Jung J, Kim J-H. A phylogenomic study of Iridaceae Juss. based on complete plastid genome sequences. Front Plant Sci. 2023;14:1066708.
47. Nanjala C, Wanga VO, Odago W, Mutinda ES, Waswa EN, Oulo MA, et al. Plastome structure of 8 *Calanthe* s.l. species (Orchidaceae): comparative genomics, phylogenetic analysis. BMC Plant Biol. 2022;22:387.
48. Dong WL, Wang RN, Zhang NY, Fan WB, Fang MF, Li ZH. Molecular evolution of chloroplast genomes of orchid species: insights into phylogenetic relationship and adaptive evolution. Int J Mol Sci. 2018;19:716.
49. Pfanzelt S, Albach DC, Von Hagen KB. Extremely low levels of chloroplast genome sequence variability in *Astelia pumila* (Asteliaceae, Asparagales). PeerJ. 2019;7:e6244.
50. Lin CS, Chen JJW, Huang YT, Chan MT, Daniell H, Chang WJ, et al. The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. Sci Rep. 2015;5:9040.
51. Lin C, Chen JJW, Chiu C, Hsiao HCW, Yang C, Jin X, et al. Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. Plant J. 2017;90:994–1006.
52. Könyves K, Bilsborrow J, Christodoulou MD, Culham A, David J. Comparative plastomics of Amaryllidaceae: inverted repeat expansion and the degradation of the *ndh* genes in *Strumaria truncata* Jacq. PeerJ. 2021;9:e12400.
53. Zheng Q, Wu Y, Ke SJ, Liu DK, Liu ZJ. Characteristics and comparative analysis of the complete plastomes of *Apostasia fujianica* and *Neuwiedia malipoensis* (Apostasioideae). Horticulturae. 2024;10:383.
54. Kim YK, Cheon SH, Hong JR, Kim KJ. Evolutionary patterns of the chloroplast genome in Vanilloid orchids (Vanilloideae, Orchidaceae). Int J Mol Sci. 2023. https://doi.org/10.3390/ijms24043808.
55. Zhou L, Chen T, Qiu X, Liu J, Guo S. Evolutionary differences in gene loss and pseudogenization among mycoheterotrophic orchids in the tribe *Vanilleae* (subfamily Vanilloideae). Front Plant Sci. 2023;14:1160446.
56. Hu C, Jiao Z, Deng X, Tu X, Lu A, Xie C, et al. The ecological adaptation of the unparalleled plastome character evolution in slipper orchids. Front Plant Sci. 2022;13:1075098.
57. Agrama HA, Tuinstra MR. Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. Afr J Biotechnol. 2003;2:334–40.
58. Li X, Zhao Y, Tu X, Li C, Zhu Y, Zhong H, et al. Comparative analysis of plastomes in Oxalidaceae: phylogenetic relationships and potential molecular markers. Plant Divers. 2021;43:281–91.
59. Mandal L. Potential applications of molecular markers in plant. Current Trends in Biomedical Engineering & Biosciences. 2018;12:85–7.
60. Jia J, Xue Q. Codon usage biases of transposable elements and host nuclear genes in *Arabidopsis thaliana* and *Oryza sativa*. Genomics Proteomics Bioinf. 2009;7:175–84.

61. Leffler EM, Bullaughey K, Matute DR, Meyer WK, Ségurel L, Venkat A, et al. Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biol. 2012;10:e1001388.
62. Dugas DV, Hernandez D, Koenen EJM, Schwarz E, Straub S, Hughes CE, et al. Mimosoid legume plastome evolution: IR expansion, tandem repeat expansions and accelerated rate of evolution in *clpP*. Sci Rep. 2015;5:16958.
63. Raubeson LA, Peery R, Chumley TW, Dziubek C, Fourcade HM, Boore JL, et al. Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. BMC Genomics. 2007;8:174.
64. Wang RJ, Cheng CL, Chang CC, Wu CL, Su TM, Chaw SM. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. BMC Evol Biol. 2008;8:36.
65. Kumar J, Choudhary K, Shelja, Anukriti, Singh H, Kumar A, et al. Molecular approaches for authentication and identification of medicinal plants. Plant Mol Biol Rep. 2025;43:1006–26.
66. Alkaraki AK, Aldmoor MA, Lahham JN, Nusair SD. DNA barcoding of selected medicinal plant species from Jordan using *matK*, *rbcL*, and *rpoC1* genes. Int J Biol Biomed Eng. 2021;15:376–411.
67. Sharma D, Maiya Shrestha T, Lamichhane J. DNA barcoding of *Myrica esculenta* (Kafal), an indigenous, multipurpose and medicinal plant species from Nepal, by *rbcL* and *matK* gene. Acta Scientific Pharmaceutical Sciences. 2020;4:11–8.
68. Yao G, Zhang YQ, Barrett C, Xue B, Bellot S, Baker WJ, et al. A plastid phylogenomic framework for the palm family (Arecaceae). BMC Biol. 2023;21:50.
69. Govaerts R. The World Checklist of Vascular Plants (WCVP). 2024.
70. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc. 2016;181:1–20.
71. Chen S, Zhou Y, Chen Y, Gu J. Fastp : an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018. https://doi.org/10.1093/bioinformatics/bty560.
72. Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, et al. Getorganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020;21:241.
73. 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;47:65–73.
74. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. 2019.
75. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33:2583–5.
76. Kurtz S. Reputer: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29:4633–42.
77. Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. 2008;9:299–306.
78. Indrabalan UB, Suresh KP, Shivamallu C, Patil SS. An extensive evaluation of codon usage pattern and bias of structural proteins p30, p54 and p72 of the African swine fever virus (ASFV). Virus Dis. 2021;32:810–22.
79. Brudno M, Malde S, Poliakov A, Do CB, Couronne O, Dubchak I, et al. Glocal alignment: finding rearrangements during alignment. Bioinformatics. 2003;19 suppl_1:54–62.
80. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34:3299–302.
81. Xiang C, Gao F, Jakovlić I, Lei H, Hu Y, Zhang H, et al. Using phylosuite for molecular phylogeny and tree-based analyses. iMeta. 2023;2:e87.
82. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, et al. Phylosuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 2020;20:348–55.
83. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
84. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFboot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35:518–22.
85. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. Modelfinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–9.
86. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-tree 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37:1530–4.

Publisher's Note