

RESEARCH

Open Access



Complete chloroplast genomes of eight *Delphinium* taxa (Ranunculaceae) endemic to Xinjiang, China: insights into genome structure, comparative analysis, and phylogenetic relationships

Chunfeng Song¹, Junwen Zhu¹ and Huimin Li^{1*}

Abstract

Background *Delphinium* L. represents a taxonomically intricate genus of significant phylogenetic and economic importance in Ranunculaceae. Despite the existence of few chloroplast genome datasets, a comprehensive understanding of genome structures and selective pressures within the genus remains unknown. Furthermore, several taxa in this genus are exclusively found in Xinjiang, China, a region renowned for its distribution and diversity of Chinese and Central Asian *Delphinium* species. Therefore, investigating the features of chloroplast genomes in this area will provide valuable insights into the evolutionary processes and phylogenetic relationships of the genus.

Results In this study, the eight newly completed chloroplast genomes are examined, ranging in length from 153,979 bp to 154,284 bp. Alongside these, analysing six previously reported taxa re-annotated in *Delphinium*, 111 unique genes are identified across all samples. Genome structure, distributions of simple sequence repeats and short dispersed repeats, as well as gene content are similar among these *Delphinium* taxa. Nine hypervariable intergenic spacers and protein coding regions, including *ndhF-trnL*^(TAG), *rpl16-intron*, *rpl33*, *rps15*, *rps18*, *trnK*^(TTT)-*trnQ*^(TTG), *trnP*^(TGG)-*psaI*, *trnT*^(GGT)-*psbD* and *ycf1*, are identified among 13 perennial *Delphinium*. Selective pressure and codon usage bias of all the plastid genes are performed within 14 *Delphinium* taxa. Phylogenetic analysis based on 14 *Delphinium* plastomes, alongside two *Aconitum* (Ranunculaceae) species serving as outgroup taxa, reveals the monophyletic nature of *Delphinium*. Our findings further discern *Delphinium* into two distinct clades: perennial species (clade I) and annual species (clade II). In addition, compared with the nrDNA ITS topology, cytological data and morphological characters, *D. mollifolium* and *D. maackianum* showed potential involvement in hybridization or polyploidization processes. Excluding these two species, the perennial *Delphinium* (clade I) exhibits a stronger consistency with the morphology-based system that utilized seed morphology.

*Correspondence:

Huimin Li
lihuimin@cnbj.net

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



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusion This study represents the first comprehensive analysis of plastomic variations among *Delphinium* taxa, based on the examination of 14 complete plastomes. The chloroplast genome structure of *Delphinium* is similar to other angiosperms and possesses the typical quadripartite structure with the conserved genome arrangement and gene features. In addition, the variation of non-coding regions is larger than coding regions of the chloroplast genome. Through DNA sequence divergence across *Delphinium* plastomes and subsequent phylogenomic analyses *ndhF-trnL*^(TAG) and *ycf1* are identified as promising molecular markers. These highly variable loci held significant potential for future phylogenetic and phylogeographic studies on *Delphinium*. Our phylogenomic analyses based on the whole plastomes, concatenation of 132 unique intergenic spacer regions, concatenation of 77 unique protein-coding genes and nrDNA ITS, all support the monophyly of *Delphinium* and perennial taxa clusters together into one clade within this genus. These findings provide crucial data for systematic, phylogenomic and evolutionary research in the genus for future studies.

Keywords China, Chloroplast genome, Comparative analysis, *Delphinium*, Phylogeny, Xinjiang

Introduction

The Genus *Delphinium* L. (Ranunculaceae) was established by Linnaeus in 1753 [1] and initially comprised only six species. Later, in 1842, De Candolle [2] introduced a classification system that divided 53 contemporaneous species into four sections, namely *D. sect. Consolida* DC., *D. sect. Delphinellum* DC., *D. sect. Staphisagria* DC. and *D. sect. Delphinastrum* DC., based on the characters of flowers, carpel and growth cycle. This classification was widely accepted by many taxonomists later [3–8]. As time passed, the perennial group *D. sect. Delphinastrum* became the largest part of the genus, comprising around 364 species [8]. However, due to the significant morphological variability and the increasing number of species, constructing clear infrageneric divisions within *D. sect. Delphinastrum* remained challenging and contentious [6–9].

Delphinium was widely distributed in the Northern Hemisphere and tropical African mountains, with approximately 500 species, more than 150 of which were native to China [6, 10–12]. Among the Chinese *Delphinium* species, the majority were perennial herbs, with only two taxa being annual [6, 11, 12]. Moreover, the taxa distributed in Xinjiang stand out as a significant distribution centre for *Delphinium* species around China and central Asia, with around 15 taxa [13, 14]. Additionally, *Delphinium* plants in China, especially in Xinjiang, had a rich history of traditional medicinal use in folk medicine, where they were used to treat various conditions such as bruises, rheumatism, toothache, and enteritis [15, 16]. They also contained chemical constituents, including flavonoids and sterols, known for their physiological activities [17]. Furthermore, some *Delphinium* species, like *D. yunnanense* Franch. and *D. grandiflorum* L., were highly valued for their ornamental qualities, highlighting the economic significance of *Delphinium* in terms of development and utilization [18, 19].

The taxonomic and phylogenetic study of *Delphinium* presented challenges within the Ranunculaceae family [20–22]. The morphological variability and large number

of species made it difficult to establish clear infrageneric divisions, especially within the perennial species that dominated the genus [8, 9, 20–22]. Although two crucial morphological characteristics like staminode color (black vs. blue) and seed morphology (seeds winged along angles vs. squamulose winged) had been used to group the perennials, their lack of correlation had led to conflicting classifications proposed by different authors [6, 8]. Traditionally, Chinese *Delphinium* species were classified into five sections based on various morphological features mainly related to the staminode color, combined with the shape of leaf, the seed morphology and the growth cycle [6]. Both infrageneric classification and taxonomic inconsistency in taxa delimitation remained a challenge in the genus [6, 8, 13, 23–27]. Furthermore, recent taxonomic revisions in different regions of China, particularly in southwestern and northwestern areas, suggested a possible decrease in the number of *Delphinium* species. For example, *D. iliense* Huth and *D. naviculare* var. *lasiocarpum* W. T. Wang were discussed here; *D. conaense* W. T. Wang was treated as a synonym of *D. bhutanicum* Munz by Yuan and Yang [24], while they were recognized as distinct species by Wang and Warnock [6], Kletter and Kriechbaum [28], respectively.

In the last two decades, molecular studies mainly focused on the phylogeny of the tribe Delphineae in Ranunculaceae [9, 29–31], with rare concentrates specifically on the infrageneric relationships within *Delphinium*, especially the Chinese group. Despite previous efforts to elucidate infrageneric relationships within the genus, several chloroplast markers shed some light on evolutionary patterns, supporting the monophyly of the *Delphinium*. However, numerous interspecies relationships based on these markers remained unresolved [9, 29–31]. Moreover, conflicting results were demonstrated in several *Delphinium* taxa between the chloroplast markers and nuclear DNA phylogenies, suggesting that unresolved intrageneric relationship might be attributed to the limited phylogenetic data available for interspecific hybridization or chloroplast capture [9, 32]. For instance,

Jabbour & Renner [29] conducted phylogenetic analyses based on three chloroplast (cp.) DNA (*trnK-matK*, *trnS-trnG*, *trnL-trnF*) and nrDNA (ITS) data, recognizing eight perennial *Delphinium* from China and North America. However, these species were divided into two geographical distribution clades with weak support. Subsequently, the authors [30] expanded their sample size to include 98 perennial *Delphinium* species from around the world, including 18 Chinese *Delphinium* taxa, and revised the phylogeny of the tribe Delphinieae (Ranunculaceae) based on nrDNA ITS region and cpDNA *trnL-trnF* data. Despite this broader dataset, the relationships among taxa remained weakly supported within the perennial *Delphinium* group. In another study, Zuo [9] investigated the evolution of seed morphology and staminode color in Chinese *Delphinium* by sampling 90 populations representing 72 perennial species. The study employed six fragment chloroplast sequences and a single-copy nuclear gene. Although the phylogenetic tree constructed from the chloroplast sequences supported the proposed classification system based on seed morphology, most interspecies relationships remained unresolved, suggesting that several species, such as *D. gyalanum* C. Marquand & Airy Shaw, *D. giraldii* Diels, *D. pulanense* W. T. Wang, experienced hybridization or chloroplast capture.

Chloroplast genomes have been frequently used in systematics due to their conserved quadripartite structure, predominantly clonal inheritance along the maternal line, and highly variable even at low taxonomic levels [33]. However, currently, only seven Chinese *Delphinium* cp. genomes were available (<https://www.ncbi.nlm.nih.gov>), which limited our knowledge of the organization and evolution of plastomes in *Delphinium* and the tribe Delphinieae [34–38].

Thus, to address these challenges in Chinese *Delphinium*, this study firstly took a comprehensive approach by analyzing the complete cp. genomes of eight Chinese *Delphinium* taxa endemic to Xinjiang. These eight cp. genomes were combined and compared with previously reported plastomes of six other Chinese *Delphinium* species [33–37]. The aims were: (1) to present the structure of cp. genome in the eight Xinjiang *Delphinium* taxa; (2) to compare the global structural patterns, investigate and screen mutational hotspots, examine variations of simple sequence repeats (SSRs) and short dispersed repeat sequences (SDRs); (3) to calculate nucleotide diversity in Chinese *Delphinium* cp. genomes for future species determination; (4) to reconstruct phylogenetic relationships among *Delphinium* species using cp. genome and nrDNA ITS region data respectively, and identify more effective molecular markers through this analysis; (5) to test for the presence of adaptive evolution in all annotated genes by analyzing selective pressure and codon

usage bias. The results of this study are expected to provide valuable information for phylogenetic and phylogeographic studies within *Delphinium* and contribute to the exploration and utilization of *Delphinium* plants.

Results

Chloroplast genome structure and features

The chloroplast genomes of the eight taxa of *Delphinium* exhibited similar structure and organization (Table 1; Figs. 1 and 2). The length of eight cp. genomes varied from 153,979 bp in *D. mollifolium* W. T. Wang to 154,284 bp in *D. shawurensense* W. T. Wang. They displayed a typical quadripartite circular structure containing a large single-copy (LSC) with lengths ranged from 84,648 bp (*D. iliense* Huth) to 85,018 bp (*D. mollifolium*), a small single-copy (SSC) with lengths varied from 16,293 bp (*D. winklerianum* Huth) to 16,342 bp (*D. shawurensense*), and a pair of inverted repeats (IRs) with lengths between 26,331 bp (*D. mollifolium*) and 26,594 bp (*D. iliense*, *D. naviculare* var. *lasiocarpum* W. T. Wang and *D. sauricum* Schischk.). The total GC content was nearly close, varied from 38.25% to 38.27%.

All the eight *Delphinium* plastomes contained the same set of 129 genes, including 84 genes encoding proteins, eight ribosomal RNAs (rRNAs) and 37 transfer RNAs (tRNAs) (Table 1; * showing the new chloroplast genomes reported in this study). Seventeen unique genes contained one (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, *rpoC1*, *trnA*^(TGC), *trnG*^(TCC), *trnI*^(GAT), *trnK*^(TTT), *trnL*^(TAA), *trnV*^(TAC)) or two (*rps12*, *clpP* and *yef3*) introns (Additional File 1: Table S1; Figs. 1 and 2). In addition, these genes could be divided into three categories according to their functions (Additional File 1: Table S1). The first type of function was mainly related to photosynthesis, with 44 unique genes; the second category of function was mainly related to cp. automatic transcription and translation, with 57 unique genes; the third category had 11 unique genes, mainly involved in other biosynthetic genes and open reading frames with unknown function.

Boundaries of IR regions, repeat structure and SSR analysis of chloroplast genomes

The potential expansions and contractions of IR borders was considered to be the main cause of cp. genome length changes and the evolutionary events in angiosperm, though relative conservation of IR/SC boundaries in plant plastomes [39, 40]. We compared the IR/SC boundaries together with the adjacent genes in the 14 *Delphinium* plastomes (including eight newly sequenced *Delphinium* cp. genomes; Fig. 3; * showing the new chloroplast genomes reported in this study). The IRa/SSC boundary was identified within *yef1* gene (with the 5' end located in the IRa region while 3' end located in the SSC region), with spanned 1060–1675 bp in the IRa region.

Table 1 Summary of characteristics of plastome sequences used in this study, including eight new chloroplast genomes of the *Delphinium* taxa, * showing the newly

Species	GenBank numbers	Total genome size (bp)	Overall GC content (%)	LSC size (bp)	IR size (bp)	SSC size (bp)	NO. total gene (unique gene)	NO. protein-coding gene (unique gene)	NO. tRNA gene (unique gene)	NO. rRNA gene (unique gene)	Reference
<i>Aconitum brachypodium</i> Diels	MT584424	155,650	38.09	86,419	26,149	16,933	129 (111)	84 (77)	37 (30)	8 (4)	[39]
<i>A. delavayi</i> Franch.	OM289058	155,733	38.08	86,362	26,229	16,913	129 (111)	84 (77)	37 (30)	8 (4)	[50]
<i>Delphinium aemulans</i> Nevski (*)	OR263583	154,245	38.27	84,809	26,561	16,314	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. anthriscifolium</i> Hance	MK253461	155,077	38.14	85,871	25,977	17,252	129 (111)	84 (77)	37 (30)	8 (4)	[35]
<i>D. brunonianum</i> Royle	NC_051554	153,926	38.30	84,512	26,559	16,296	129 (111)	84 (77)	37 (30)	8 (4)	[36]
<i>D. candelabrum</i> var. <i>monanthum</i> (Hand.-Mazz.) W. T. Wang	MW246165	153,995	38.25	84,862	26,543	16,047	129 (111)	84 (77)	37 (30)	8 (4)	[36]
<i>D. ceratophorum</i> Franch.	MK253460	154,245	38.27	84,801	26,560	16,324	129 (111)	84 (77)	39 (31)	8 (4)	[35]
<i>D. elatum</i> var. <i>sericeum</i> W. T. Wang (*)	OR263584	154,219	38.28	84,780	26,561	16,317	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. iliense</i> Huth (*)	OR263586	154,161	38.27	84,648	26,594	16,325	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. maackianum</i> Regel	NC_047293	154,484	38.13	85,055	26,564	16,301	129 (111)	84 (77)	37 (30)	8 (4)	[37]
<i>D. mollifolium</i> W. T. Wang (*)	OR263588	153,979	38.25	85,018	26,331	16,299	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. naviculare</i> var. <i>lasiocarpum</i> W. T. Wang (*)	OR263587	154,171	38.27	84,656	26,594	16,327	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. sauricum</i> Schischk. (*)	OR263590	154,255	38.25	84,765	26,594	16,302	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. shawurensense</i> W. T. Wang (*)	OR263585	154,284	38.26	84,820	26,561	16,342	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. winklerianum</i> Huth (*)	OR263589	154,235	38.26	84,754	26,594	16,293	129 (111)	84 (77)	37 (30)	8 (4)	This study
<i>D. yunnanense</i> Franch.	MW246156	154,053	38.26	84,639	26,551	16,312	129 (111)	84 (77)	37 (30)	8 (4)	[36]

Similarly, the IRb/SSC boundary was located within *ycf1* gene (with the 5' end located in the IRb region while 3' end in the SSC) and *ndhF* gene (with the 5' end located in the SSC region while 3' end in the IRb), with the former expanded 29–30 bp in the SSC region and the latter expanded 31–32 bp in the IRb region. However, in the case of *D. anthriscifolium* Hance sample (MK253461), the IRb/SSC boundary only located within *ycf1* gene, with an expansion length of 11 bp in the SSC region. The IRb/LSC boundary exhibited obviously varied. Four samples, including *D. ceratophorum* Franch., *D. iliense*, *D. naviculare* var. *lasiocarpum* and *D. winklerianum*, had the boundary located within the *rps19* gene, with an expansion length of 1–34 bp. The remaining 11 samples

were either 0–5 bp away from the IRb/LSC boundary, except for the *D. mollifolium* sample (OR263588), where the boundary was located in *rpl2* gene (with the 5' end located in the IRb region while 3' end in the LSC) with an expansion length of 164 bp.

We detected six types of SSR (Additional File 2: Table S2; Fig. 4A) for each species in 14 *Delphinium* cp. genome, in which the number of total SSRs was from 51 (*D. anthriscifolium*) to 57 [*D. brunonianum* Royle, *D. candelabrum* var. *monanthum* (Hand.-Mazz.) W. T. Wang, *D. elatum* var. *sericeum* W. T. Wang, *D. mollifolium*]. Most cp. genome SSRs, with the proportion from 73.58% (*D. shawurensense*) to 79.25% (*D. naviculare* var. *lasiocarpum*) out of the total number of SSRs, were distributed in

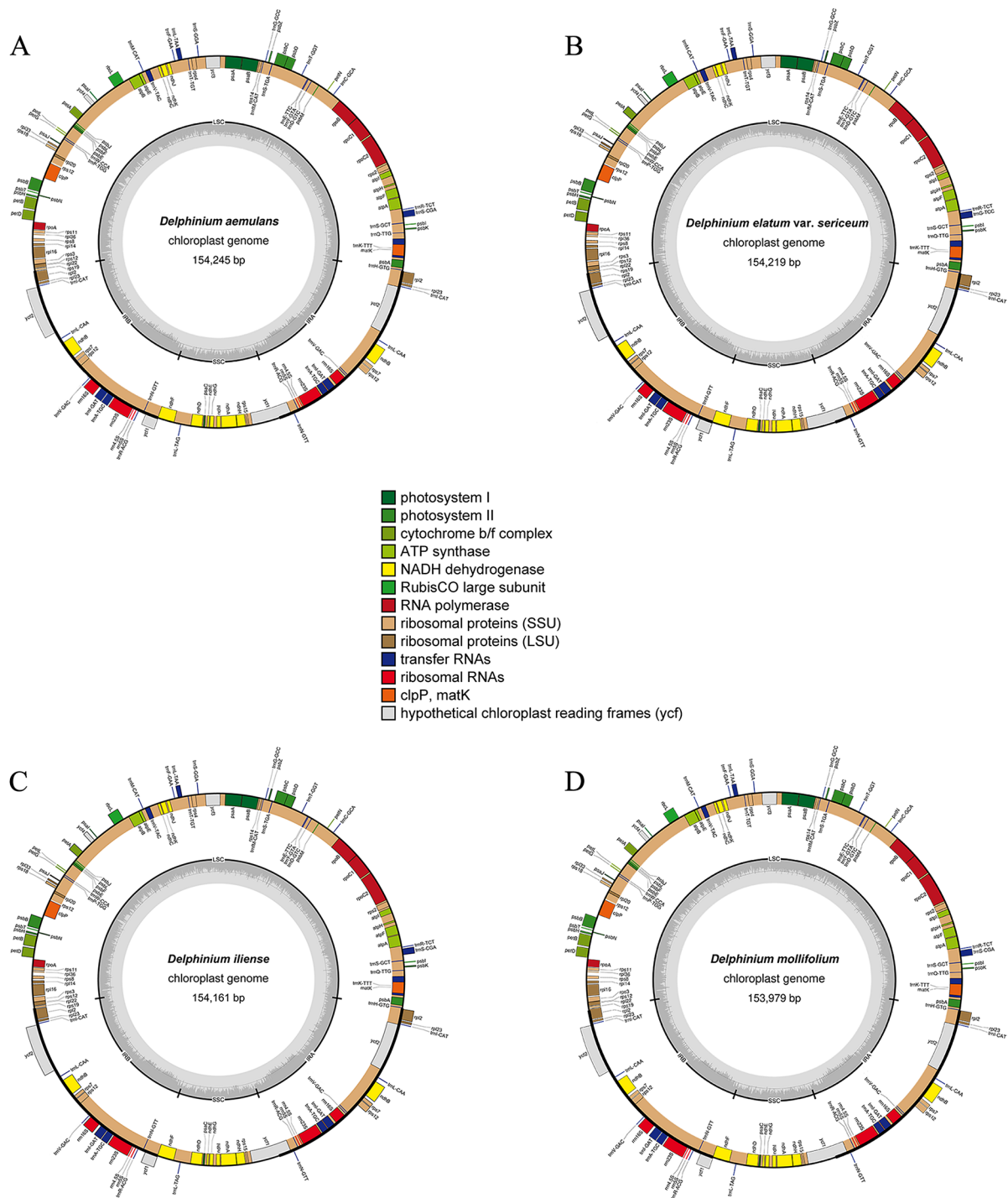


Fig. 1 Plastomes of four *Delphinium* taxa, including *D. aemulans*, *D. elatum* var. *sericeum*, *D. ilense* and *D. mollifolium*. The outer circle shows the genes at each locus, and inverted repeat regions are indicated with thicker lines. Genes on the outside of the outer circle are transcribed in a counterclockwise direction, while genes on the inside of the outer circle are transcribed in a clockwise direction. The inner circle indicates the range of the large single-copy (LSC), small single-copy (SSC), and the inverted repeats (IRs), and also shows a GC content graph of the genome. In the GC content graph, the dark gray lines indicate GC content, while light gray lines indicate the AT content at each locus

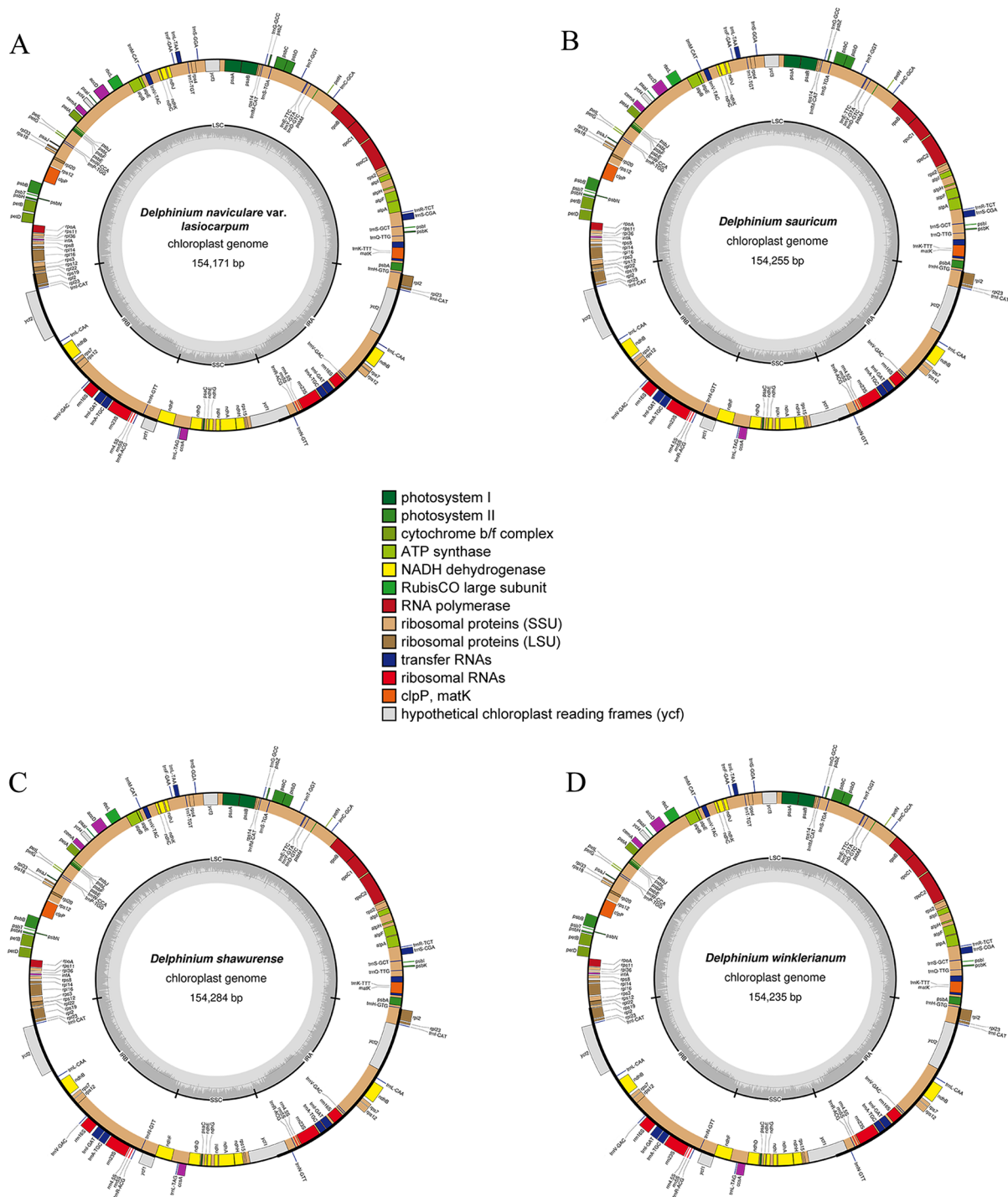


Fig. 2 Plastomes of four *Delphinium* taxa, including *D. naviculare* var. *lasiocarpum*, *D. sauricum*, *D. shawurense* and *D. winklerianum*. The outer circle shows the genes at each locus, and inverted repeat regions are indicated with thicker lines. Genes on the outside of the outer circle are transcribed in a counter-clockwise direction, while genes on the inside of the outer circle are transcribed in a clockwise direction. The inner circle indicates the range of the large single-copy (LSC), small single-copy (SSC), and the inverted repeats (IRs), and also shows a GC content graph of the genome. In the GC content graph, the dark gray lines indicate GC content, while light gray lines indicate the AT content at each locus



Fig. 3 Comparison of LSC, inverted-repeats (IRs), and SSC junction positions among 14 *Delphinium* plastomes, * showing the new chloroplast genomes reported in this study

the LSC regions. The SSRs distributed in the SSC region ranged from 14.55% (*D. aemulans* Nevski) to 19.61% (*D. anthriscifolium*) and in the IR regions varied from 3.51% (*D. mollifolium* and *D. winklerianum*) to 10.53% (*D. candelabrum* var. *monanthum*) (Fig. 4B). Among these SSRs, the mono-nucleotide A/C/G/T repeat units occupied the

highest proportion with 86.27–96.36%, and the di-nucleotide repeats (AT/TA) and tri-nucleotide repeats (AAT/ATA) units accounted for 1.81–13.73% and 0–1.92% out of the total number of SSRs, respectively (Fig. 4C).

Meanwhile, more than 30 bp of base repeats in all samples and distinct forms of these long repeats, including

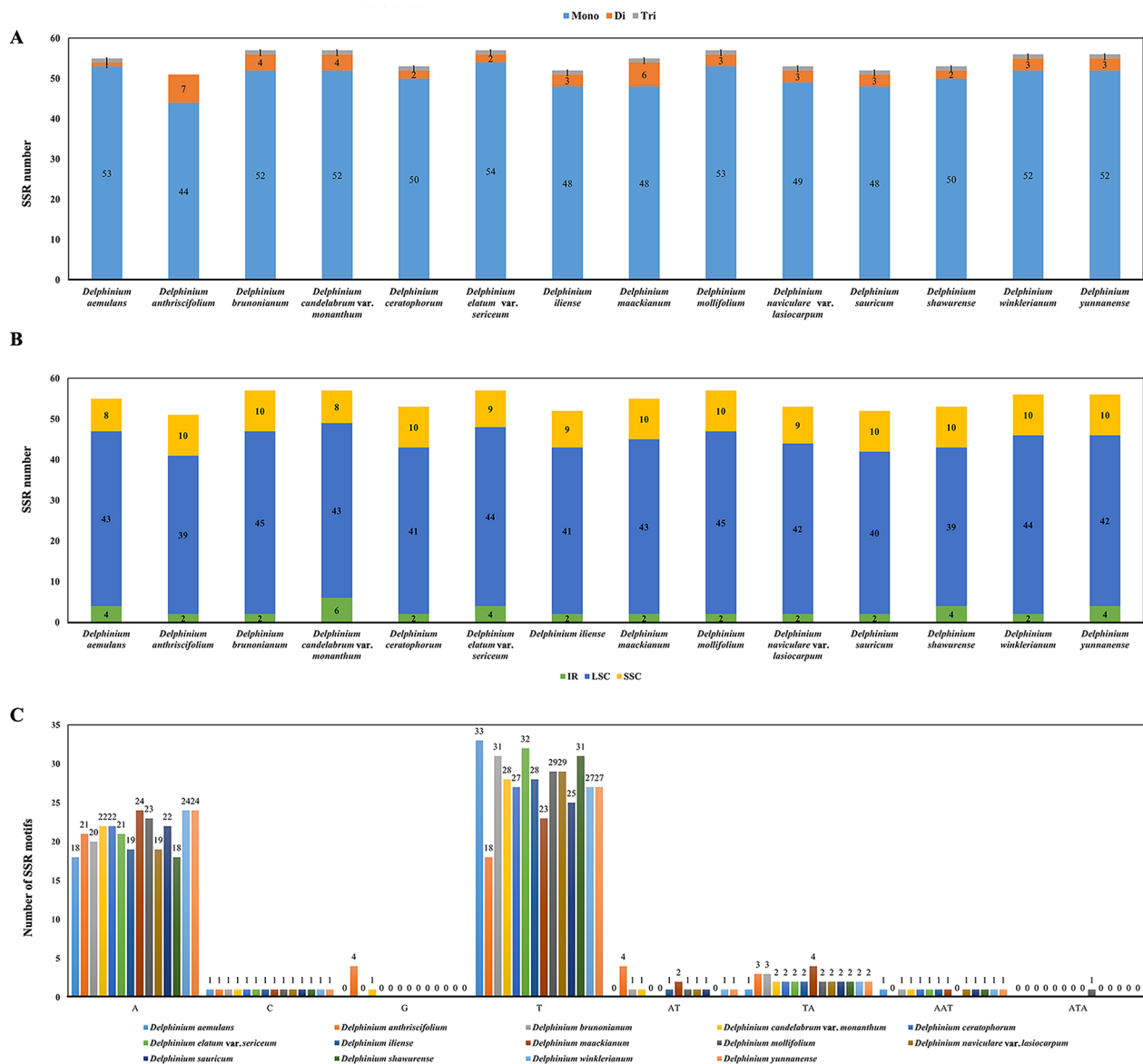


Fig. 4 Statistics of SSRs in 14 *Delphinium* taxa samples. (A) Number of SSRs types. (B) Number of SSRs distributed in LSC, SSC and IR regions. (C) Distribution of different types and sizes of SSRs motifs in the plastid genomes

complementary, forward, palindromic and reversed, were also analyzed (Additional File 3: Table S3; Fig. 5). For these 14 *Delphinium* cp. genomes, the size of the top three most frequently long repeats were 30 bp, 31 bp and 42 bp. The distribution of repeats per genome, and length of repeat and number of such repeated sequences per species were shown in Fig. 5A, respectively. In each taxon, the number of long repeats ranged from 18 (*D. aemulans*) to 28 (*D. anthriscifolium*); and the number of complementary, forward, palindromic and reversed repeats were 0–1, 6–11, 12–15 and 0–3, respectively (Fig. 5B). Most long repeats were distributed in intergenic

areas, and a few in shared genes or introns, such as *ycf2* and *ycf3*-intron.

Genome comparison, hypervariable regions

The value of nucleotide variability (Pi) value among 14 *Delphinium* cp. genomes revealed that the intergenic spacer (IGS) regions were more variable than gene regions (Additional File 4: Table S4; Fig. 6). The SSC regions showed the highest average nucleotide diversity (Pi=0.00998), followed by the LSC (Pi=0.006619) and IR (Pi=0.001231) regions. 32 hypervariable sites in LSC region with Pi≥0.01 were screened (Fig. 6), namely *accD-psaI*, *atpH-atpI*, *cemA*, *matK*, *ndhC-trnV^(TAC)*,

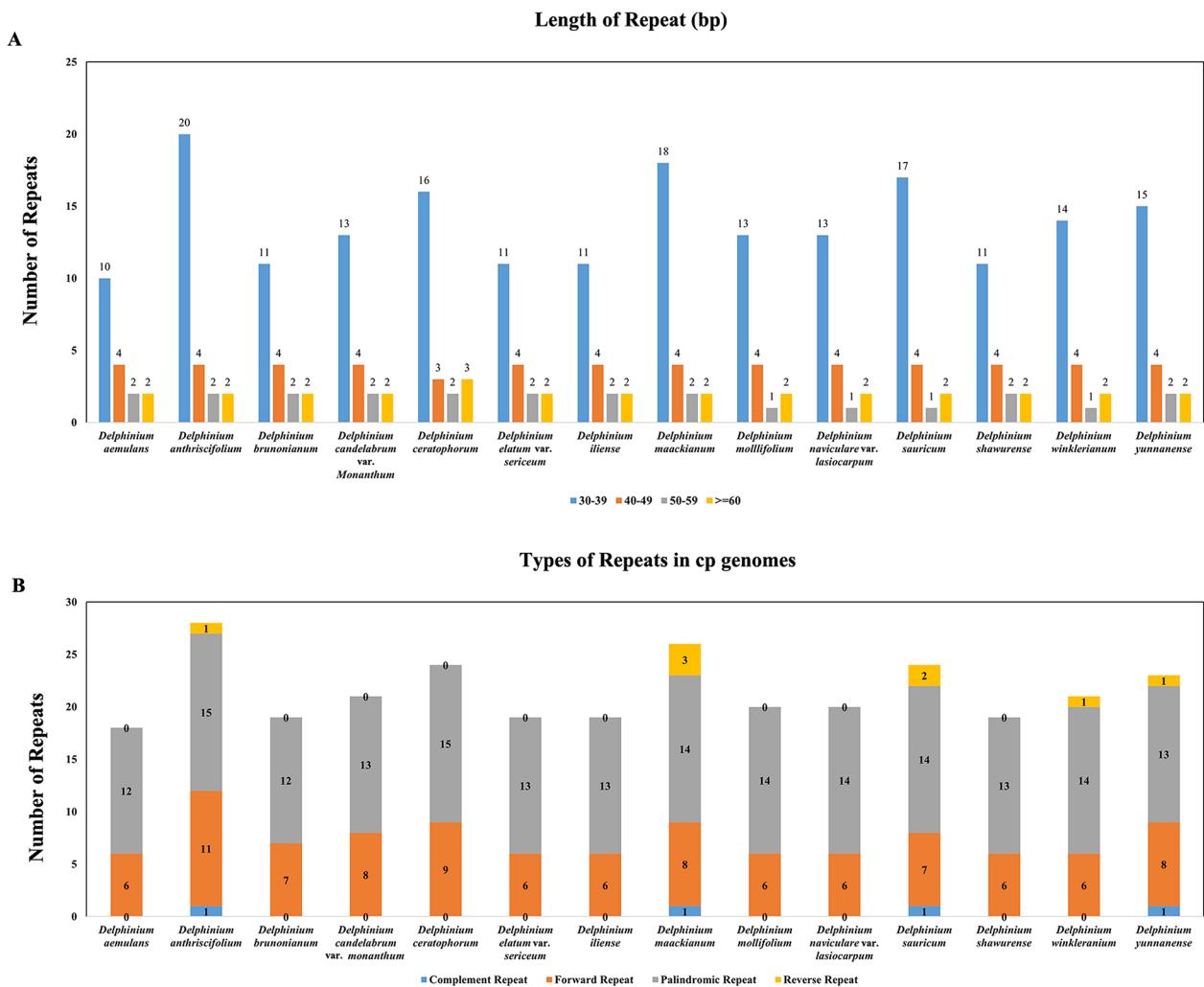


Fig. 5 Statistics of repeats in 14 *Delphinium* taxa samples. (A) Number of different lengths of repeats. (B) Number of four types of repeats

petN-psbM, *psaI*, *psaJ*, *psaJ-rpl33*, *psbE-petL*, *psbM-trnD*^(GTC), *rpl16-intron*, *rpl20*, *rpl33-rps18*, *rpoB-trnC*^(GCA), *rps3*, *rps8*, *rps18*, *rps18-rpl20*, *trnC*^(GCA)-*petN*, *trnD*^(GTC), *trnE*^(TTC)-*trnT*^(GGT), *trnF*^(GAA)-*ndhJ*, *trnG*^(TCC), *trnK*^(TTT)-intron, *trnK*^(TTT)-*trnQ*^(TTG), *trnP*^(TGG)-*psaJ*, *trnS*^(GCT)-*trnG*^(TCC), *trnT*^(GGT)-*psbD*, *trnT*^(TGT)-*trnL*^(TAA), *ycf3-trnS*^(GGA), *ycf4-cemA*; and eight hypervariable sites with $P_i \geq 0.01$ in SSC regions were also screened in Fig. 6, namely *ccsA*, *ccsA-ndhD*, *ndhD*, *ndhF*, *ndhF-trnL*^(TAG), *rps15*, *rps15-ycf1* and *ycf1*.

Meanwhile, considering the annual species, *Delphinium anthriscifolium*, was definitely different from perennial groups. The value of nucleotide variability (P_i) value among 13 perennial *Delphinium* cp. genomes was also conducted. The results showed that the IR regions were observed to have lower P_i value than LSC and SSC regions. The SSC regions showed the highest average nucleotide diversity ($P_i=0.005164$), followed by the LSC ($P_i=0.003326$) and IR ($P_i=0.000519$) regions. Six

hypervariable sites in LSC having $P_i \geq 0.01$, were *rpl16-intron*, *rpl33*, *rps18*, *trnK*^(TTT)-*trnQ*^(TTG), *trnP*^(TGG)-*psaJ*, *trnT*^(GGT)-*psbD*; while three hypervariable sites in SSC regions with $P_i \geq 0.01$, namely *ndhF-trnL*^(TAG), *rps15* and *ycf1*. (Additional File 5: Table S5; Fig. 7).

The mVISTA results showed that the non-coding regions were more variable than the coding regions, the LSC and SSC regions had higher levels of sequence divergence than the two IR regions, and the IGS regions were the most divergent regions (Fig. 8). The highly divergent regions among 14 chloroplast genomes occurred six in the IGS regions, four in the LSC regions, including *rbcL-accD*, *rpoB-trnC*^(GCA), *trnT*^(GGT)-*psbD*, *trnP*^(TGG)-*psaJ*, and two near the boundary between IRa and SSC region: *ccsA-ndhD*, *ndhF-trnL*^(TAG). Apart from these regions, one region *ycf1* also showed high sequence variation (Fig. 8).

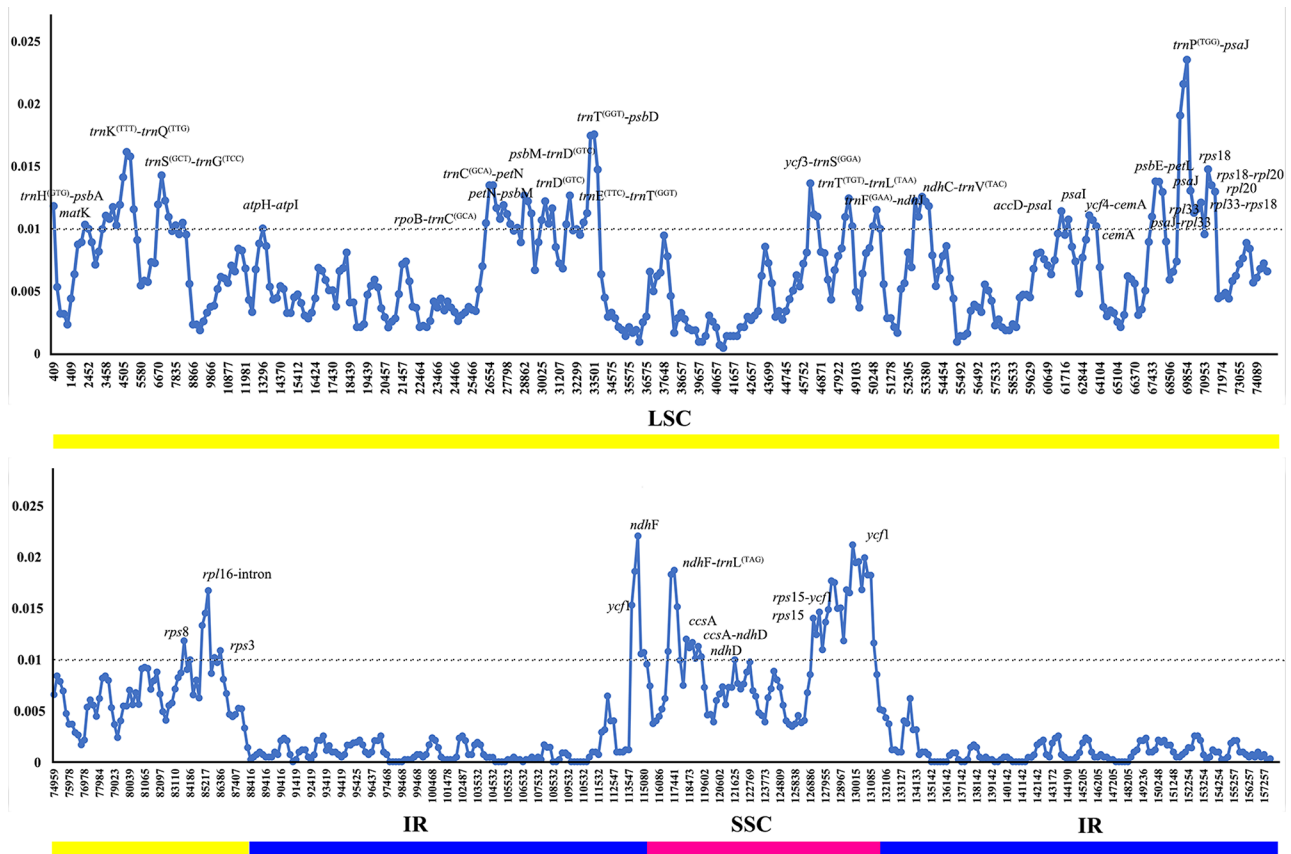


Fig. 6 Comparison of nucleotide variability values (Pi) among 14 *Delphinium* chloroplast genomes. The x-axis indicates the position of the midpoint of a window, while the y-axis represents the nucleotide diversity of each window. The different colored lines at the bottom indicate the locations of these genes in various regions of the plastomes

Codon usage analysis

We detected the synonymous codon usage of 77 unique PCGs in the 14 *Delphinium* and calculated several related parameters, including the effective number codons (ENC), codon bias index (CBI) and relative synonymous codon usage (RSCU). The ENC and CBI of all these 77 unique PCGs varied a wide range, with the former ranging from 24.824 to 61 and the latter ranging from 0.268 to 0.853 (Additional File 6: Table S6). The results showed that these genes were expressed in different levels probably due to the frequency of optimal codons [41]. The PCGs contained a total of 22,525 to 22,556 codons in the 14 *Delphinium* plastomes, including stop codons. Leucine (Leu; 2328–2339) was the most abundant amino acid, while Cysteine (Cys; 252–263) showed the least abundance in the cp. genome of these taxa (Additional File 7: Table S7).

The RSCU value analysis showed that almost all amino acids were encoded by one to six synonymous codons, except methionine and tryptophan (Met and Trp; RSCU=1). Almost Half of these codons (32/61; not including stop codon) had RSCU ≥ 1, in which most (29/32) ended with the base A or U. Meanwhile, about

half of codons (29/61; not including stop codon) had RSCU < 1, in which majority (27/29) ended with the base C or G. All three stop codons were present, with UAA being the most frequently used among these 77 unique PCGs in the 14 *Delphinium* (Additional File 7: Table S7; Fig. 9).

Selective pressure

The ω ratio (d_N/d_S) of 77 unique PCGs among these 16 species in Ranunculaceae were calculated to estimate the selective pressure. A total of six genes (*clpP*, *petN*, *psbJ*, *psbZ*, *rpl23* and *ycf1*) in m0 model were found to be under positive selection (ω ratio > 1), while 35 genes (*atpF*, *atpH*, *ccsA*, *cemA*, *clpP*, *infA*, *ndhA*, *ndhE*, *ndhG*, *ndhH*, *ndhK*, *petA*, *petD*, *petN*, *psaA*, *psaC*, *psaI*, *psbE*, *psbF*, *psbI*, *psbJ*, *psbZ*, *rpl14*, *rpl20*, *rpl22*, *rpl23*, *rps3*, *rps4*, *rps7*, *rps8*, *rps15*, *rps19*, *ycf1*, *ycf2*, *ycf4*) in m2 model were identified as being under positive selection. The value of ω ratio was significantly different ($P < 0.05$) among these taxa for two genes (*psaA* and *rpl20*) based on likelihood ratio tests (LRTs) (Additional File 8: Table S8).

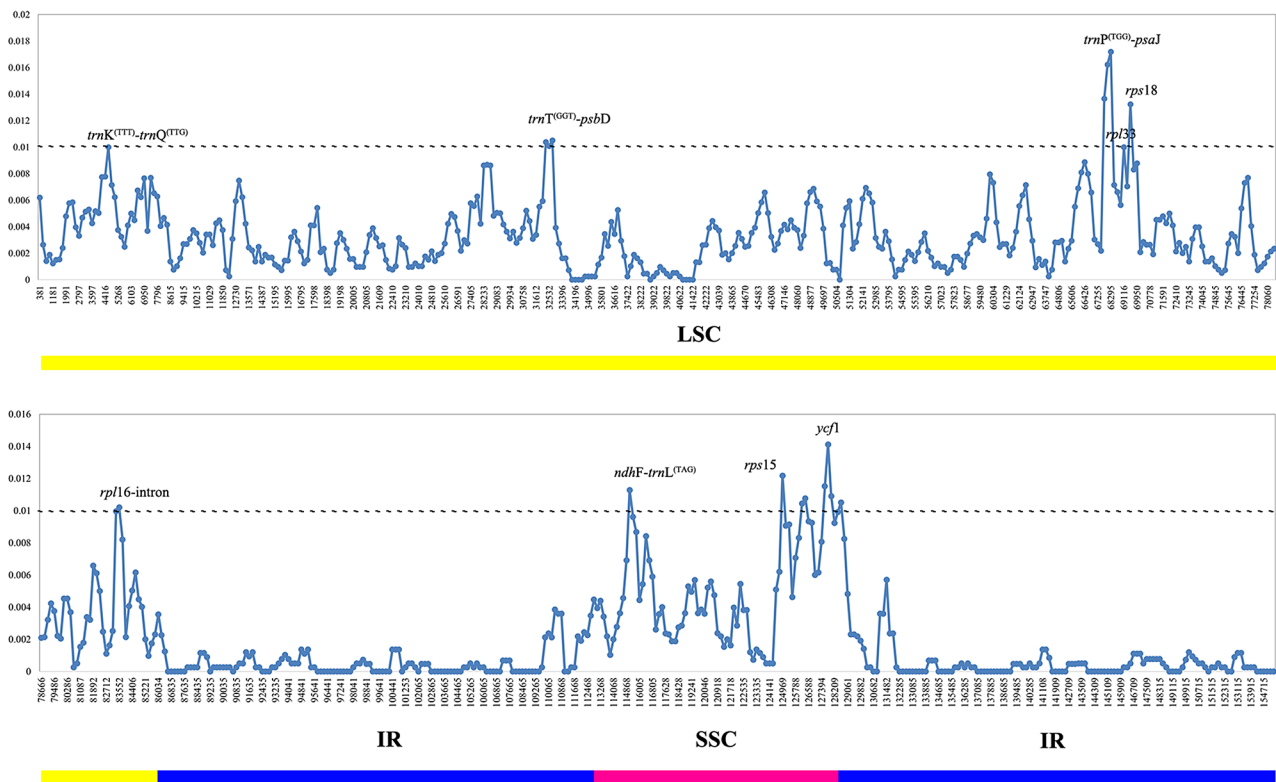


Fig. 7 Comparison of nucleotide variability values (Π) among 13 perennial *Delphinium* chloroplast genomes. The x-axis indicates the position of the mid-point of a window, while the y-axis represents the nucleotide diversity of each window. The different colored lines at the bottom indicate the locations of these genes in various regions of the plastomes

Phylogenetic analysis

We used three datasets, including the whole complete plastid genome sequences, concatenation of 132 unique IGS regions, and concatenation of 77 unique PCGs regions to construct the phylogenetic relationships among the 14 *Delphinium* species, respectively, with *Aconitum brachypodum* Diels and *A. delavayi* Franch. as outgroups by using ML method. Despite minor discrepancies, the results of these three topologies were found in high congruence (Fig. 10).

Our results showed that the genus *Delphinium* was monophyletic with strong support and contained two major clades which were fully supported [Bootstrap Support (BS)=100] as sister groups: clade I and II (Fig. 10). The clade I comprised all perennial *Delphinium* samples divided into two strongly supported subclades (I–1 and I–2). Subclade I–1 (BS=100) contained eight *Delphinium* taxa all collected from Xinjiang, China. However, only slightly different supporting values were observed at nodes based on the different sequence datasets. For instance, the nodes (red showing in Fig. 10) in subclade I–1 derived from the dataset of the whole complete plastid genome showed stronger supports (BS=100 and BS=100; Fig. 10A) than those derived from the concatenation of 132 unique IGS regions (BS=95.7 and

BS=100; Fig. 10B) and the concatenation of 77 unique PCGs (BS=97.6 and BS=100; Fig. 10C). Additionally, the supporting value (red showing in Fig. 10) of subclade I–2 derived from the whole complete plastid genome (BS=98.5; Fig. 10A) and 77 unique PCGs (BS=89.5; Fig. 10C) was stronger than it derived from the IGS regions (BS=85.1; Fig. 10B), while the supporting values of one nodes (blue showing in Fig. 10) in subclade I–2 derived from the whole cp. genome (BS=98.5; Fig. 10A) were stronger than it derived from the concatenation of 132 unique IGS regions (BS=87.7; Fig. 10B) and concatenation of 77 unique PCGs regions (BS=94.9; Fig. 10C). Besides, the resolution of previously used three sequence fragments, including *rbcL*, *trnS*^(TGA)-*trnG*^(TCC) and *trnL*^(CAA) [29–31] also concatenated here was also evaluated for *Delphinium* species, which showed in poorly supporting value (Fig. 11A).

A total of 12 hypervariable or high Π value regions (*accD-rbcL*, *ccsA-ndhD*, *ndhF-trnL*^(TAG), *rpoB-trnC*^(GCA), *rpl16*, *rpl33*, *rps15*, *rps18*, *trnK*^(TTT)-*trnQ*^(TTG), *trnP*^(TGG)-*psaI*, *trnT*^(GGT)-*psbD* and *ycf1*) and concatenation of these 12 regions were also evaluated for phylogenetic analysis in our study (Figs. 11B–I and 12A–E). Moreover, the analysis of sequence alignments revealed the presence of parsimony-informative characters across multiple

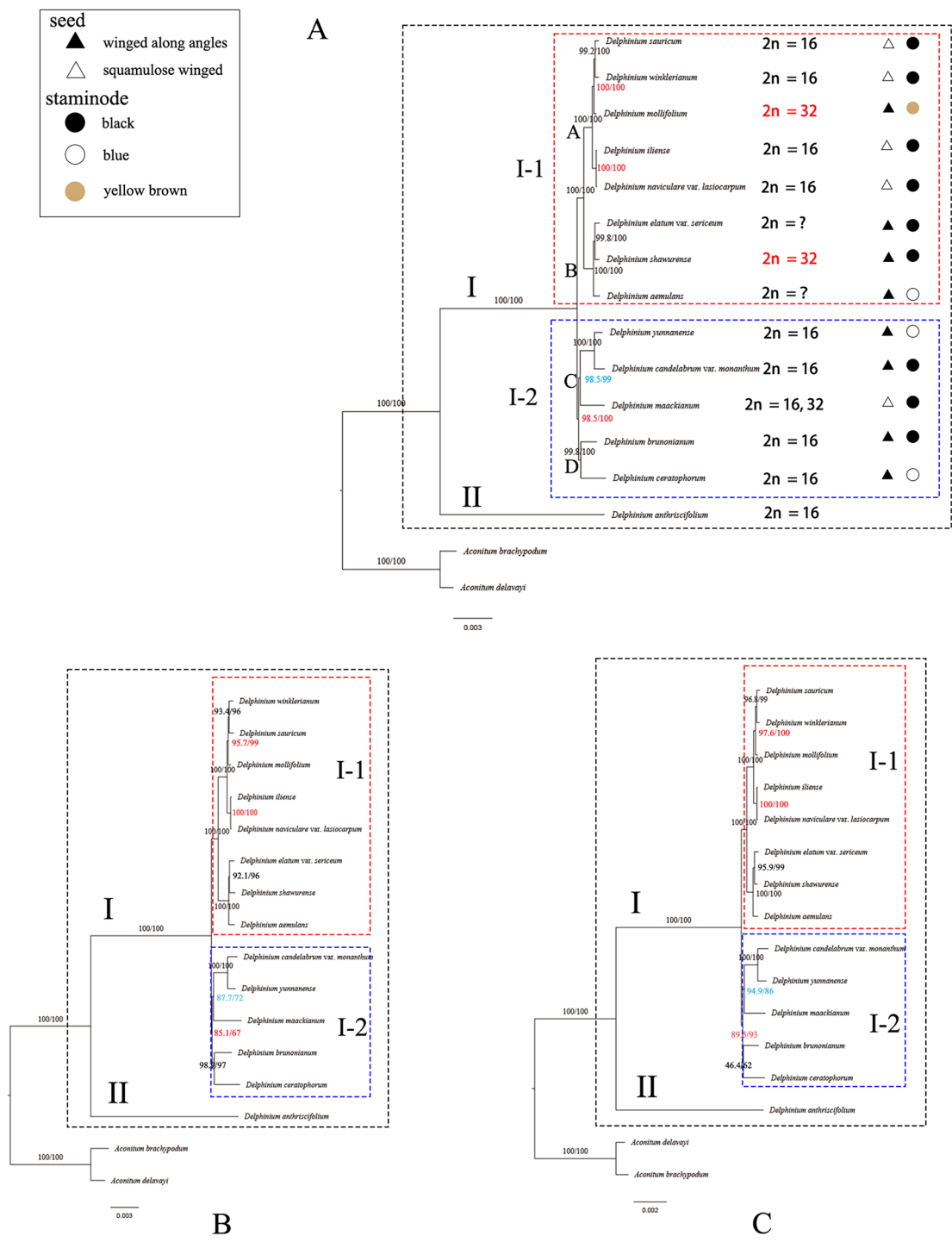


Fig. 10 Phylogenetic trees based on complete cp. genomes (A), concatenation of 132 unique IGS region (B), and concatenation of 77 unique PCGs regions (C) resulting from the maximum likelihood (ML) analysis of 14 *Delphinium* samples and two *Aconitum* species as outgroups. The bootstrap support values in ML analysis are displayed at nodes

(Fig. 10A), two fragment sequences topological trees (Figs. 11D and E and 12D) based on *ndhF-trnL*^(TAG) and *ycf1* performed well in dividing the perennial *Delphinium* into two groups as the whole cp. genome. In addition, the concatenation of 12 hypervariable or high Pi

value regions (Fig. 12E) yielded highly similar topological results to the whole cp. genome, with different supporting values. For example, the nodes in clade I-2 derived from the dataset of a concatenation of 12 regions showed strong supports (BS=86.6; red showing in Fig. 12E) lower

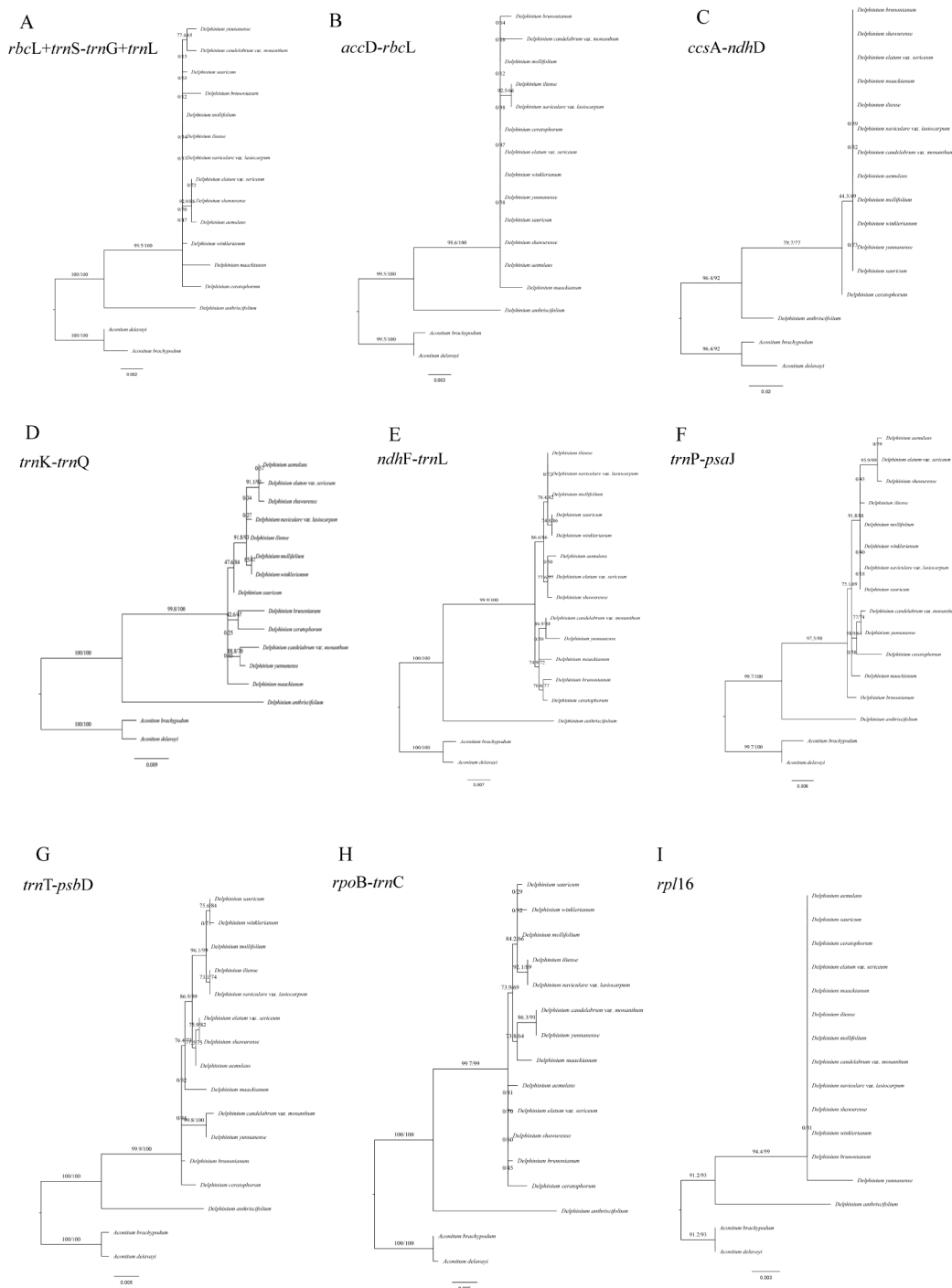


Fig. 11 Phylogenetic trees among 14 *Delphinium* samples and two *Aconitum* species as references inferred from maximum likelihood (ML) analysis. **(A)** concatenation of *rbcL*, *trnS*^(TGA)-*trnG*^(TCC) and *trnL*^(CAA). **(B)** *accD-rbcL*. **(C)** *ccsA-ndhD*. **(D)** *trnK*^(TTT)-*trnQ*^(TTG). **(E)** *ndhF-trnL*^(TAG). **(F)** *trnP*^(TGG)-*psaJ*. **(G)** *trnT*^(GGT)-*psbD*. **(H)** *rpoB-trnC*^(GCA). **(I)** *rpl16*

than those from the whole cp. genome (BS=98.5; red showing in Fig. 10A), concatenation of 77 unique PCGs regions (BS=89.5; red showing in Fig. 10C).

Furthermore, the relationship was also reconstructed by using ML analysis based on nrDNA ITS herein among nine *Delphinium* and two *Aconitum* taxa (Fig. 12F). The

nrITS topology was similar to the topologies inferred from the plastid genome sequences with different placement in some species. For instance, in contrast to the tree derived from plastid genome sequences, in nrITS tree, *D. aemulans* and *D. elatum* var. *sericeum* formed a strongly supported clade sister to *D. shawurensense* with strongly

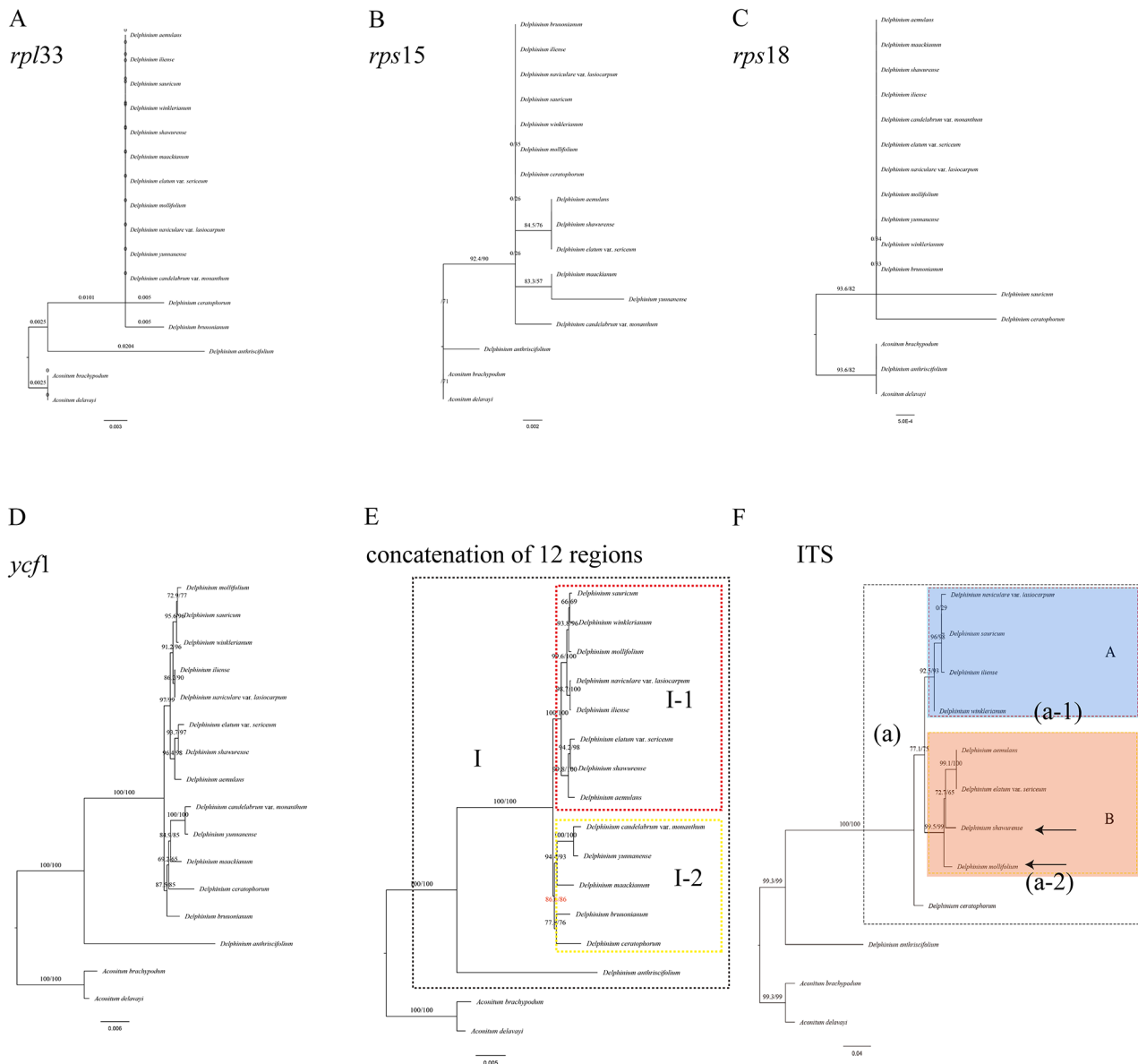


Fig. 12 Phylogenetic trees based on *rpl33* (A), *rps15* (B), *rps18* (C), *ycf1* (D) and concatenation of 12 regions (E) among 14 *Delphinium* samples and two *Aconitum* species as references; nrDNA ITS region (F) among 10 *Delphinium* samples and two *Aconitum* species as references inferred from maximum likelihood (ML) analysis. Note: the arrowed showing the tetraploid taxa, *D. mollifolium* and *D. shawurensense*

supported (BS=72.7; Fig. 12F) and nested within one branch with *D. mollifolium* in clade a–2.

Discussion

Plastome features in *Delphinium*

In this study, the cp. genomic structure, gene order and GC content among eight newly sequenced *Delphinium* taxa were highly conserved and nearly similar, which were also identical to other cp. genomes in angiosperms [42]. The size of the 14 *Delphinium* plastomes varied from 153,926 bp (*D. brunonianum*; NC_051554) to 155,077 bp (*D. anthriscifolium*; MK253461) (Table 1). The *Delphinium* cp. genome sequenced here, all contained

a total of 129 genes (including 111 unique genes), with the total GC content ranging from 38.13% to 38.30% (Table 1). However, compared with the previously published plastomes of other seven taxa in *Delphinium* [35–37], some taxa were found to contain different numbers of genes in different samples, for instance, *D. anthriscifolium* (MK253461), *D. brunonianum* (NC_051554), *D. candelabrum* var. *monanthum* (MW246165), *D. ceratophorum* (MK253460), *D. maackianum* Regel (NC_047293) and *D. yunnanense* (MW246156) were reported contained 128, 131, 130, 128, 129 and 130 genes, respectively, whereas all annotated here contained only 129 genes. To eliminate the influences of references

and annotation software used, the 14 samples were re-annotated using Plastid Genome Annotator (PGA) and Geneious Prime 2023.01.1, with *Nicotiana tabacum* L. (NC001879) and *Aconitum brachypodium* (MT584424) as the reference genome. Additionally, the tRNA genes were verified by tRNA-SE and ARAGORN. Unexpectedly, upon examining all the 14 sequences re-annotated, we found that only 129 genes and no gene loss were observed in this study (Table S1). Moreover, it should be noted that the plastome of *D. grandiflorum* (NC_049872) previously reported, which had been associated with ambiguous or incorrect information and potential misidentifications, were not included in our analysis.

The contraction and expansion of IR/SC boundary regions were usually considered as an important process involved in cp. genome variation within Ranunculaceae [35] and other angiosperm plastomes [43]. Furthermore, this phenomenon had proven to be particularly informative for evolutionary studies within specific groups [44]. However, minor variations were observed in the IR/SC boundary regions with no significant expansion or contractions among the 14 *Delphinium* plastomes (Fig. 3). The length of the IR region ranged from 25,977 bp to 26,594 bp. Only the *rpl2* gene with an expansion length of 164 bp for *D. mollifolium* expanded to the LSC region; the remaining 13 *Delphinium* samples were entirely located within the IRb region. And the *ndhF* gene with contraction length of 101 bp away from the IRb region only in *D. anthriscifolium* (MK253461). These results were also similar to the contraction or expansion event in the cp. genome of other genera in Ranunculaceae, such as *Gymnaconitum* (Stapf) Wei Wang & Z. D. Chen, *Nigella* L., *Beesia* Balf. f. & W. W. Sm., *Actaea* L., *Souliea* Franch. and *Ranunculus* L. [35, 45, 46].

Repeats and SSRs were widely analyzed in plant plastomes [47]. The variations of SSRs in cp. genomes were widely used to analysis the genome structure and diversity in population or species levels [48, 49]. Our findings indicated that mono-nucleotide repeats were the predominant type of repeat in the cp. genomes (Fig. 4), consistent with previous studies conducted in angiosperm cp. genomes [45, 50]. Among various types of SSRs, mono-nucleotide SSRs composed of A/T repeats exhibited higher abundance in the cp. genome. This observation aligned with prior reports suggesting that the prevalence of A/T repeats might be attributed to the relatively easier conversion of A/T compared to C/G in the plant cp. genome [51]. SSRs had also demonstrated their values in assessing genetic diversity within and between populations, as well as in studying the parentage of hybrid individuals in specific groups [52, 53]. Repeats variation in type, number, and location in different taxa, which were used to identify mutational hotspots and phylogenetic relationships [46, 54]. Four

types of repeats (complement repeats, forward repeats, palindromic repeats and reverse repeats) were detected, among which palindromic repeats were the most common type of repeats (Fig. 5). Moreover, the number and variety of repeats in *D. anthriscifolium*, *D. maackianum*, *D. sauricum*, and *D. yunnanense* were found to exceed those present in other taxa within the genus *Delphinium*. To determine whether these repeats and SSRs could be effectively employed for phylogenetic analysis of the *Delphinium* genus, additional investigations will be required in the future.

Potential molecular markers in *Delphinium*

The plastid genomes exhibited a high degree of conservation in terms of genetic replication mechanisms and uniparental inheritance, while displaying relatively high levels of genetic variation due to low selective pressure [55]. Consequently, the cp. genome has proven successful in resolving relationships within the Ranunculaceae family [41, 45, 46] and other angiosperms [50, 56]. Our phylogenetic analysis, based on complete cp. genomes, as well as the concatenation of 132 unique IGS regions and the concatenation of 77 unique PCGs (Fig. 10), consistently supported a well-defined clade (clade I) for perennial *Delphinium*, aligning with previous studies [9, 29–31]. Notably, the phylogenetic trees based on complete plastid genomes (Fig. 10A) exhibited stronger support compared to other concatenation sequences (Fig. 10B, C). However, when using three fragments of *rbcL*, *trnS*^(TGA)-*trnG*^(TCC), and *trnL*^(CAA) as done in previous studies [9, 29, 30], our results indicated low resolution among the 14 *Delphinium* taxa (Fig. 11A).

Furthermore, our findings revealed that non-coding regions exhibited higher divergence compared to coding regions, consistent with patterns observed in numerous angiosperms [57]. Variable chloroplast sequences were widely used for phylogenetic analysis and taxonomic identification [58]. Therefore, we constructed 13 maximum likelihood (ML) trees using a total of 12 hypervariable or high Pi value regions and their concatenation (Figs. 11B–I and 12A–E). These trees demonstrated that only two fragment sequences (Figs. 11D and E and 12D) having higher number of parsimony informative characters, namely *ndhF-trnL*^(TAG) and *ycf1*, performed well in differentiating perennial *Delphinium* into two groups, similar to the whole cp. genome, except for the highly consistent concatenation topology (Fig. 12E). It should be noted, however, that the low resolution in the ML tree based on *rpl16* (Fig. 11I) and *rpl33* (Fig. 12A) may be attributed to the loss of the *rpl32* gene in the tribe Delphinieae [37, 59], which leads to compensatory changes in the plastid-encoded *rpl* subunits, rendering them unreliable markers for phylogeny in *Delphinium*. Thus, the highly variable sequences generated in this

study, especially *ndhF-trnL*^(TAG) and *ycf1*, represented promising potential molecular markers for phylogeny reconstruction and DNA barcoding identification in *Delphinium* plants.

Positive selection among genes

It is noted that all genes are basically subjected to a certain degree for natural selection, and the highly expressed genes might be selected by the evolutionary forces [41]. Codon usage analysis played a crucial role in unravelling the evolutionary process, genome structure and selection pressure on genes [60]. In this study, the remarkable similarities observed in RSCU among 14 *Delphinium* taxa strongly suggested the presence of common environmental factors influencing their evolutionary trajectory. Additionally, a notable bias towards a lower frequency in base C or G at the third codon position, as compared to A or U, was observed (Fig. 9). These findings were consistent with previous investigations conducted on cp. genomes of other angiosperm [51, 61], lending further support to our conclusions. The degeneracy of genetic code enabled the expression of genetic variation within a gene, resulting in the production of diverse proteins across different species [62]. Meanwhile, we conducted an analysis of different codon usage frequencies on 77 unique PCGs across 16 taxa under positive pressure (ω ratio > 1). The results revealed an upper limit of ω ratio = 999, indicating a lack of synonymous substitutions along the concerned branch. This phenomenon, occurring for various reasons, requires further investigation in the future [63, 64]. However, the likelihood ratio test (LRT) value remained valid in our analysis. Our results indicated that only two genes were subject to significantly strong positive selective pressure. Specifically, one plastid gene associated with photosystem I (*psaA*) exhibited significantly strong positive selective pressure ($P < 0.005$ based on likelihood ratio tests) across two branches, with a relatively lower Codon Bias Index (CBI) value ($< \text{CBI median} = 0.505$). Additionally, another plastid gene related to ribosomal protein (*rpl20*) also showed significantly strong positive selective pressure ($P < 0.005$ based on likelihood ratio tests) across three *Delphinium* species, displaying a higher CBI ($> \text{CBI median} = 0.505$). The differential codon usage bias observed in these genes suggested varying frequencies of rare and optimal codons, potentially influencing their expression patterns and functional evolution. Furthermore, the differences between selective pressure and codon usage frequencies among these plastid genes implied potential functional divergence among the 14 *Delphinium* taxa.

Phylogenetic relationships

In our study, we conducted a phylogenetic analysis of Chinese *Delphinium* using the entire plastid genome

(Fig. 10A) and nrDNA ITS (Fig. 12F) to investigate the monophyly, infrageneric classification and assess their status. Previous research on the systematics of Chinese *Delphinium* had primarily relied on several molecular fragments [9, 29, 30] and morphological investigations [6, 8]. However, our phylogenetic trees showed inconsistencies with all the previously reported molecular phylogenetic studies. For example, Zuo [9] found that *D. elatum* var. *sericeum* (endemic to Xinjiang) formed a clade with other Xinjiang species in the cpDNA tree but was segregated in the nuclear gene tree. In contrast, our results demonstrated that the samples from Xinjiang, China, including *D. elatum* var. *sericeum*, formed a well-supported clade (clade I–1) in the plastid genome tree (Fig. 10A) and a well-supported clade (clade a) in the nrDNA ITS tree (Fig. 12F).

Analyzing the plastid topology (Fig. 10A), we observed that clade I comprised all perennial *Delphinium* taxa, while clade I–1 exclusively included samples collected from Xinjiang, China. Clade I–1 further divided into two well-supported clades, clade A and clade B. Interestingly, the status of taxa in clade A and clade B were mostly consistent with the morphology-based system based on seed morphology with a slight difference [8]. For instance, within clade A, *D. mollifolium* exhibited brown staminodes and winged seeds along angles, deviating from other taxa characterized by black staminodes and squamulose winged seeds. Similarly, in clade B, while all taxa displayed winged seeds along angles, *D. aemulans* stood out with blue staminodes as opposed to the predominant black staminodes observed in other taxa. These observations echo the findings of Zuo [9], highlighting occasional inconsistencies between molecular and morphology-based taxonomic systems. Furthermore, our analysis of the nrDNA ITS tree (Fig. 12F) revealed that all Xinjiang samples clustered within clade a, which further bifurcated into two well-supported clades, clade a–1 and clade a–2. Intriguingly, distinct seed morphology traits delineated the taxa within these clades, with clade a–1 exhibiting squamulose winging and clade a–2 displaying winging along angles. Unlike the chloroplast-based phylogenetic tree proposed by Zuo [9], which supported the morphology-based classification system primarily based on seed morphology, our nrDNA ITS tree demonstrated greater consistency with this classification scheme. Moreover, our study underscores the importance of incorporating additional samples and molecular fragments from nuclear ribosomal DNA for more comprehensive taxonomic investigations in the future.

Additionally, previous cytology research [9, 13, 65–69] reported that *D. mollifolium* and *D. shawurense* were tetraploid with a chromosome number of $2n = 32$, while all other taxa were diploid ($2n = 16$), except for *D. aemulans* and *D. elatum* var. *sericeum*, for which chromosome

number remained unknown (Additional File 10, 11: Table S10; Fig. S1). Morphologically, the staminode color in *D. mollifolium* was yellow brown, representing an intermediate character between the two crucial colors of black and blue (Fig. 10A). Moreover, the placement of *D. mollifolium* exhibited discrepancies between the plastid topology and nrDNA tree. Given the occurrence of hybridization and chloroplast capture events in Ranunculaceae, as noted by many authors [9, 70, 71], resolving the conflicting status of *D. mollifolium* would necessitate the incorporation of more nrDNA markers and samples.

Moreover, Wang and Yang [14] reported that Xinjiang province gathered the numerous basal taxa of evolutionary branches in Chinese *Delphinium* as well as taxa from Central Asia. They [14] also suggested that this area clearly represented the densest population of basal *Delphinium* species in China and even Central Asia. It encompassed both evolutionary early branching lineages and relatively basal components, along with a few more evolutionarily late branching lineages, representing low-level, middle-level, and high-level species in the phylogeny of *Delphinium*. Therefore, to accurately ascertain the status of clade I–1 and the significance of the Xinjiang groups in the context of *Delphinium* species, it was crucial to include plastomes from additional Chinese *Delphinium* samples or Central Asia.

Continuing with the analysis of the plastid topology (Fig. 10A), the samples within clade I–2 were divided into two well-supported clades: clade C and clade D. In clade C, *D. yunnanense* and *D. candelabrum* var. *monanthum*, characterized with seed winging along angles, different staminode colors, and *D. maackianum*, which with two types of chromosome numbers, characterized with squamulose winged seeds, and black staminodes, formed a fully supported clade. In clade D, *D. brunonianum* and *D. ceratophorum*, with winged seeds along angles and different staminode colors, clustered together in a well-supported clade.

As a result, excluding *D. mollifolium* and *D. maackianum*, two species with abnormal chromosome numbers, the perennial *Delphinium* (clade I) exhibited greater consistency with the morphology-based system that utilized seed morphology [8]. To further investigate the relationships among these species, especially *D. mollifolium* and *D. maackianum*, which might involve hybridization or polyploidization, future studies should increase the sample size and incorporate additional nrDNA markers. By expanding the sample size and utilizing more molecular markers, we will better understand the genetic relationships within these species. This approach will enable a comprehensive analysis of potential hybridization events or chromosomal changes, providing insights into the underlying mechanisms influencing the observed characteristics in *D. mollifolium* and *D. maackianum*.

Due to the high morphological variability, particularly in staminode color and seed morphology, two important but uncorrelated characters, taxonomic inconsistencies persist in the delimitation of taxa within the genus *Delphinium* [7, 8, 27]. For example, based on our observations of living plants in the field and examination of herbarium specimens, including type material, we previously demonstrated that *D. iliense* exhibited high variability in the indumentum of peduncles, pedicels, bracteoles, sepals, and carpels, as well as in the shape and position of bracteoles on pedicels. Consequently, we redefined this species and synonymized two names, including *D. iliense* var. *angustatum* Huth and *D. naviculare* var. *naviculare*, with it. However, we found that *D. naviculare* var. *lasiocarpum*, much like *D. iliense*, displayed high morphological variability. Although Borodina-Grabovskaya [72] synonymizing *D. naviculare* var. *lasiocarpum* with *D. naviculare* var. *naviculare* (a synonym of *D. iliense*), our plastid tree placed these two taxa together in a fully supported clade (Fig. 10A), while they revealed less proximity in the nrDNA tree (Fig. 12F). Moreover, through examination of herbarium specimens and living plants, we distinguished them based on the indumentum of stems, peduncles, and pedicels. A detailed investigation of the identity of *D. naviculare* var. *lasiocarpum* will be presented separately.

Conclusion

This study represents the first comprehensive analysis of plastomic variations among *Delphinium* taxa, based on the examination of 14 complete plastomes. The chloroplast genome structure of *Delphinium* is similar to other angiosperms and possesses the typical quadripartite structure with the conserved genome arrangement and gene features. However, their size varies owing to the expansion/contraction of IR/SC boundaries. The variation of non-coding regions is larger than coding regions of the chloroplast genome. DNA sequence divergence across *Delphinium* plastomes and phylogenomic analyses reveal that *ndhF-trnL*^(TAG) and *ycf1* are promising molecular markers. Therefore, these highly variable loci should be valuable for future phylogenetic and phylogeographic studies on *Delphinium*. Our phylogenomic analyses based on the whole plastomes, concatenation of 132 unique IGS regions, concatenation of 77 unique PCGs sequences and nrDNA ITS sequence, all support the monophyly of *Delphinium* and perennial taxa clusters together into one clade within this genus. These results will provide important data for systematic, phylogenomic and evolutionary research in the genus for future studies.

Materials and methods

Sampling, DNA extraction, chloroplast genome sequencing, assembling, and annotation

Plant materials of the eight *Delphinium* taxa were collected in the field during 2022 from Xinjiang Province in China. Fresh leaves were sampled and dried in silica gel immediately. Voucher specimens were deposited in the herbarium of Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (NAS) and collection information were listed in the Additional File 12: Table S11. In addition, six complete chloroplast genomes of *Delphinium* species (Table 1) and two of *Aconitum* species (Table 1) that publicly available in NCBI GenBank were downloaded with annotations. Total genomic DNA was following a modified cetyltrimethylammonium bromide CTAB method [73]. DNA integrity was examined by electrophoresis in 1% (w/v) agarose gel, and concentration was measured by Qubit® DNA Assay Kit in Qubit® 3.0 Fluorometer (Invitrogen, USA).

High-quality DNA libraries were constructed by shearing the genomic DNA into short fragments with approximately 350 bp before sequenced on Illumina platform and generated 150 bp paired-end reads at Novogene Bioinformatics Technology Co., Ltd. (Tianjin, China). Genomes assembly were performed using the GetOrganelle pipeline [74–76] based on the sequenced clean data. Bandage v.5.6.0 [77] was used to visualize and manually correct the assembly results. The annotation of the chloroplast genomes was performed in PGA [78]. Further annotation confirmation was compared with four sequences in the same tribe Delphinieae, *Aconitum brachypodium* (MT584424), *A. delavayi* (OM289058), *Delphinium anthriscifolium* (MK253461), *D. ceratophorum* (MK253460). Manual correction of start/stop codons and intron/exon boundaries was performed in Geneious Prime 2023.0.1 [79]. All transfer RNA (tRNA) genes were proofread with the web server tRNAscan-SE 2.0 (<http://lowelab.ucsc.edu/tRNAscan-SE/>) [80] and ARAGORN 1.2.38 (<http://www.trna.se/ARAGORN/>) [81]. All genome maps were drawn by Organellar Genome DRAW (OGDRAW) (<http://ogdraw.mpimg.de/>) [82]. The complete cp. genome sequences and gene annotation of the eight newly assembled *Delphinium* taxa samples were deposited in GenBank (Table S1). Meanwhile, all the six cp. genomes in *Delphinium* reported previously were re-annotated.

Genome comparison, codon usage analyses, plastid genomic variations and sequences repeat analysis

Using MAFFT v7.490 [83] to align the total 14 cp. genomes sequences (Table S1) for examining the divergence regions among *Delphinium* species. The aligned sequences were performed in Shuffle-LAGAN model via mVISTA program (<http://genome.lbl.gov/vista/mvista/>

[submit.shtml](#)) with the annotated cp. genome sequence of *D. aemulans* (GenBank accession no. OR263583) as a reference genome. DnaSP v6 [84] was applied to examine the sequence divergence hotspots with conducting a sliding window analysis to calculate pi values among the cp. genomes, with windows size of 600 bp and step size of 200 bp.

IRscope software was used for the 14 *Delphinium* cp. genome sequences to visualize their IR/SC boundaries. CodonW [85] was used to analyze codon usage bias for all PCGs in the *Delphinium* plastome. Parameters such as ENC, CBI, and RSCU were calculated. ENC and CBI evaluated codon bias at the gene level, while RSCU observed and expected codon frequencies [86, 87]. Amino acid (AA) frequency was determined as the percentage of codons encoding the same AA out of the total codons. The program DnaSP v6 [84] was used for examination and complementary analysis of the codon usage bias results obtained from CodonW software [855].

SSRs were identified by Web-based simple sequence repeats finder MISA-web (<https://webblastipk-gaterleben.de/misa/>), with minimum numbers of 10 repeat units for mono-, 6 repeat units for di-, 5 repeat units for tri-, tetra-, penta-, and hexa-nucleotide SSRs. The maximum length of a sequence between two SSRs was set as 10. REPuter was implemented to detect the short dispersed repeats [88], including forward, reverse, complement and palindromic, with the following parameters: a maximal repeat size of 5000, a minimal repeat size of 30, and hamming distance of 3.

Phylogenetic analysis

A total of 16 complete cp. genome sequences were used for phylogenetic analysis, including eight newly and six previous reported *Delphinium* taxa, as well as *Aconitum brachypodium* and *A. delavayi* in Ranunculaceae as outgroups [20, 37] in this study (Table 1). Phylogenetic analyses were performed using ML method in the IQ-tree program [89] with auto substitution model and 1000 bootstrap replicates for evaluating the node support. FigTree v1.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) was used to visualize the resulting trees. The analyses were carried out based on the following 18 datasets, including the complete plastid DNA, concatenation of 132 unique IGS regions, concatenation of 77 unique PCGs, the concatenation of two IGS regions and one gene (including *rbcL*, *trnS*^(TGA)-*trnG*^(TCC) and *trnL*^(CAA)) that previously studied in the tribe Delphinieae [29–31], 12 high pi value or hypervariable regions (*accD-rbcL*, *ccsA-ndhD*, *ndhF-trnL*^(TAG), *rpoB-trnC*, *rpl16*, *rpl33*, *rps15*, *rps18*, *trnK*^(TTT)-*trnQ*^(TTG), *trnP*^(TGG)-*psaI*, *trnT*^(GGT)-*psbD* and *ycf1*), concatenation of these 12 regions and the nrDNA ITS region.

Selective pressure analysis

Selective pressures were examined throughout the phylogenetic tree of *Delphinium* for 77 unique PCGs. The Easy-CodeML software [90] in PAML v4 [91] was used to assess the nonsynonymous (d_N) and synonymous (d_S) substitution rates of each plastid gene. The ω ratio (d_N/d_S) indicated the selection pressure on genes, ω less than 1 revealed purification selection, equal to 1 revealed neutral evolution, and greater than 1 revealed positive selection [92]. We tested different hypotheses via branch models, M0: the one-ratio model (m0) assumed the same d_N/d_S ratio (ω ratio) for all branches in the phylogeny; M2: the two-ratio model (m2) assumed the outgroup branch had ω ratio that differed from that throughout the rest of the tree [91]. LRTs were used to perform pairwise comparisons of these models [93].

Abbreviations

AA	Amino acid
bs	Bootstrap support
cp	Chloroplast
CBI	Codon bias index
CTAB	Cetyl trimethylammonium bromide
d_N	Nonsynonymous
d_S	Synonymous
ENC	Effective number codons
IGS	Intergenic spacers
IRs	Inverted repeats
ITS	Internal transcribed spacer of ribosomalDNA
LRTs	Likelihood ratio tests
LSC	Large singlecopy
ML	Maximumlikelihood
NCBI	National Center for Biotechnology
PCGs	Protein-coding genes
PGA	Plastid genome annotator
Pi	Nucleotide diversity/polymorphism
rRNA	Ribosomal RNA
RSCU	Relative synonymous codon usage
SDR	Short dispersed repeats
SSC	Small singlecopy
SSR	Simple sequence repeat
tRNA	Transfer RNA

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-024-05279-y>.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6
Supplementary Material 7
Supplementary Material 8
Supplementary Material 9
Supplementary Material 10
Supplementary Material 11

Supplementary Material 12

Acknowledgements

The authors thank Jiang Xu, Lan Wu and Xue-Jia Zhang, who helped to collect and plant the materials used for the experiments; and also grateful for Qiang Lai, Ming-Song Wu for kind advice on the data analysis software.

Author contributions

S-CF: Conceptualization, Methodology, Validation, Software, Reviewing and Editing the manuscript; Z-JW: Software, Analyses, and Reviewing; L-HM: Investigation, Conceptualization, Methodology, Analyses, Writing the original draft and Editing the manuscript. All authors have read and agree to the published version of the manuscript.

Funding

This research was funded by National Natural Science Foundation of China, grant number 32370220; Natural Science Foundation of Jiangsu Province, grant number BK20200294; Forestry Administration of Jiangsu Province, grant number LYKJ[2022]02.

Data availability

The datasets presented in this study can be found in online repositories. The names of the repository and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/genbank/>, OR263583, OR263584, OR263585, OR263586, OR263587, OR263588, OR263589 and OR263590.

Declarations

Ethics approval and consent to participate

All the samples used in this study were collected in accordance with the applicable national and local regulations. The plant specimens gathered were not on the list of nationally protected plants, nor were they obtained from a national park or nature reserve. No specific permission was required for their collection in line with the national and local legislation at the time of collection. The molecular experiments conducted were in compliance with the relevant laws of China. All the voucher specimens were identified by Huimin Li.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (Nanjing Botanical Garden Mem. Sun Yat-Sen), Nanjing 210014, Jiangsu, China

Received: 26 July 2023 / Accepted: 10 June 2024

Published online: 26 June 2024

References

1. Linnaeus C. *Species Plantarum*. Laurentii Salvii: Holmiae; 1753.
2. de Candolle AP. *Delphinium* L. Prodrômus systematis naturalis regni vegetabilis, sive, enumeratio contracta ordinum generum specierumque plantarum huc usque cognitarium, juxta methodi naturalis, normas digesta. Parisii: Sumptibus Sociorum Treuttel et Würtz; 1824. pp. 51–56.
3. Franchet A. Exposition synoptique et description des *Delphinium* de la Chine. In: Bulletin de la Société Philomathique de Paris. 1893. pp. 157–187.
4. Tamura M. Morphology, phylogeny and ecology of the Ranunculaceae VI. *Sci Rep Osaka Univ*. 1966;15:13–35.
5. Tamura M. Ranunculaceae. In: Hiepko P, editor. Die natürlichen Pflanzenfamilien. Berlin: Duncker und Humblot; 1995. pp. 223–496.

6. Wang WT, Warnock MJ. *Delphinium* Linnaeus. In: Wu ZY, Raven PH, Hong DY, editors. *Flora of China*. & St. Louis: Science Press & Missouri Botanical Garden; 2001. pp. 223–74. Beijing.
7. Wang WT. *Delphinium* L. In: Anonymous, editor. *Flora reipublicae popularis sinicae*. Beijing: Science; 1979. pp. 326–462.
8. Malyutin NI. The system of the genus *Delphinium* (Ranunculaceae). *Botanicheskiy Zhurnal*. 2001;86:120–30.
9. Zuo L. Evolution of the seed morphology and staminode colour in the perennial taxa of *Delphinium* (Ranunculaceae) and a re-evaluation of their taxonomic applicability: evidence from chloroplast and nuclear DNA sequence data. For the degree of Master of Science. South China Botanical Garden, Chinese Academy of Sciences; 2013.
10. Espinosa F, Deroin T, Malécot V, Wang W, Pinedo M, Nadot S, et al. Historical note on the taxonomy of the genus *Delphinium* L. (Ranunculaceae) with an amended description of its floral morphology. *Adan*. 2021;43:9–18.
11. Gan Q, Li X. A New species of *Delphinium* (Ranunculaceae) from Hubei, China. *Novo*. 2017;25:430–432.
12. Wang WT. A revision of the genus *Delphinium* (Ranunculaceae) of China (I). *Guihaia*. 2019;39:1425–1469.
13. Li HM. Taxonomic revision of *Delphinium* L. Ranunculaceae) from Xinjiang, China. Doctor of philosophy in Botany. University of Chinese Academy of Sciences; 2019.
14. Wang B, Yang C. Study on the Genus *Delphinium* L. from Xinjiang. *Journal August 1st Agri Coll*. 1995;18.
15. Chen S, Meng L, El-Demerdash FM, Zhou L, Rizvi SAH, Cui L, et al. Review of compounds and Pharmacological effects of *Delphinium*. *J Chem*. 2020;2020:e9375619.
16. Zhang JF. Studies on Diterpenoid Alkaloids and Biological Activity of Five Specific Traditional Chinese Medicines Cao Wu in Xinjiang. Doctor Degree. Southwest Jiaotong University; 2018.
17. Lin C-Z, Zhao Z-X, Xie S-M, Mao J-H, Zhu C-C, Li X-H, et al. Diterpenoid alkaloids and flavonoids from *Delphinium Trichophorum*. *Phytochemistry*. 2014;97:88–95.
18. Guan W, Li S. Ornamental germplasm resources of genus *Delphinium* in Yunnan. *Subtropical Plant Sci*. 2002;31:61–64.
19. Cheng Y, Guan W, Li F. Characterisation and evaluation of interspecific cross-incompatibility in two *Delphinium* species. *Folia Horticulturae*. 2020;32:307–317.
20. Cossard G, Sannier J, Sauquet H, Damerval C, de Craene LR, Jabbour F, et al. Subfamilial and tribal relationships of Ranunculaceae: evidence from eight molecular markers. *Plant Syst Evol*. 2016;302:419–431.
21. Hadidchi A, Attar F, Ullah F. Using microscopic techniques for taxonomic implications of seed and fruits of *Delphinium* L. (Sensu lato) (Ranunculaceae). *Microsc Res Tech*. 2020;83:99–117.
22. Kashin A, Kritskaya T, Bogoslov A, Shilova I, Parkhomenko A, Ishmuratova M et al. Taxonomic revision of *Delphinium* (Ranunculaceae) in the South-East of European Russia. *Plant Syst Evol*. 2021;307.
23. Yuan Q, Yang Q-E. Taxonomic studies on the genus *Delphinium* (Ranunculaceae) from China (XXII): clarifying morphological distinction between *D. Drepanocentrum* and *D. Umbrosum* and synonymizing *D. umbrosoides* with *D. Drepanocentrum*. *Phytotaxa*. 2022;572:243–258.
24. Yuan Q, Yang Q-E. Taxonomic studies on the genus *Delphinium* (Ranunculaceae) from China (XXIII): reduction of *D. conaense* to the synonymy of *D. Bhu-tanicum*, a species newly reported for China. *Phytotaxa*. 2023;599:291–300.
25. Yuan Q, Yang Q. Taxonomic studies on the genus *Delphinium* (Ranunculaceae) from China (XIX): the identity of *D. Calthifolium* from northern Sichuan. *Phytotaxa*. 2019;404:255–260.
26. Li HM, Yuan Q, Yang Q-E. Taxonomic studies on the genus *Delphinium* (Ranunculaceae) from China (XVII): five new synonyms of *D. afghanicum*, a newly recorded species for China. *Phytotaxa*. 2019;393:89.
27. Li HM, Yuan Q, Yang Q-E. Taxonomic studies on the genus *Delphinium* (Ranunculaceae) from China (XVIII): towards clarification of the confusion of *D. iliense* with special reference to observations on living plants in the Ili region in northwestern Xinjiang. *Phytotaxa*. 2019;403:1.
28. Kletter C, Kriechbaum M. *Tibetan Medicinal plants*. CRC; 2001.
29. Jabbour F, Renner SS. *Consolida* and *Aconitella* are an annual clade of *Delphinium* (Ranunculaceae) that diversified in the Mediterranean basin and the Irano-Turanian region. *TAXON*. 2011;60:1029–1040.
30. Jabbour F, Renner SS. A phylogeny of Delphinieae (Ranunculaceae) shows that *Aconitum* is nested within *Delphinium* and that late miocene transitions to long life cycles in the Himalayas and Southwest China coincide with bursts in diversification. *Mol Phylogenet Evol*. 2012;62:928–942.
31. Xiang K-L, Aytaç Z, Liu Y, Espinosa F, Jabbour F, Byng JW, et al. Recircumscription of *Delphinium* subg. *Delphinium* (Ranunculaceae) and implications for its biogeography. *Taxon*. 2017;66:554–566.
32. Koontz JA, Soltis PS, Soltis DE. Using phylogeny reconstruction to test hypotheses of hybrid origin in *Delphinium* section *Diedropetala* (Ranunculaceae). *Syst Bot*. 2004;29:345–357.
33. Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol*. 2011;76:273–297.
34. Duan H, Lu Y, Duan X, Zhou X, Wang C, Tian F et al. Characterization of the complete chloroplast genome of *Delphinium grandiflorum* L. *Mitochondrial DNA Part B*. 2019.
35. He J, Yao M, Lyu R-D, Lin L-L, Liu H-J, Pei L-Y, et al. Structural variation of the complete chloroplast genome and plastid phylogenomics of the genus *Asteropyrum* (Ranunculaceae). *Sci Rep*. 2019;9:15285.
36. Li Q, Guo X, Yuan F, Nima C, Dongzhi D, Duojie, et al. Characterization of the complete chloroplast genome of the musk Larkspur *Delphinium brunonianum* (Ranunculaceae). *Mitochondrial DNA Part B*. 2020;5:2394–2396.
37. Park S, An B, Park S. Recurrent gene duplication in the angiosperm tribe Delphinieae (Ranunculaceae) inferred from intracellular gene transfer events and heteroplasmic mutations in the plastid *matK* gene. *Sci Rep*. 2020;10.
38. Yuan Y. *Delphinium candelabrum* var. *monanthum* chloroplast, complete genome. 2020.
39. Xi J, Lv S, Zhang W, Zhang J, Wang K, Guo H et al. Comparative plastomes of *Carya* species provide new insights into the plastomes evolution and maternal phylogeny of the genus. *Front Plant Sci*. 2022;13.
40. Zong D, Qiao Z, Zhou J, Li P, Gan P, Ren M, et al. Chloroplast genome sequence of triploid *Toxicodendron vernicifluum* and comparative analyses with other lacquer chloroplast genomes. *BMC Genomics*. 2023;24:56.
41. Xiang K-L, Mao W, Peng H-W, Erst AS, Yang Y-X, He W-C et al. Organization, phylogenetic marker exploitation, and Gene Evolution in the Plastome of *Thalictrum* (Ranunculaceae). *Front Plant Sci*. 2022;13.
42. Yu T, Gao J, Liao P-C, Li J-Q, Ma W-B. Insights into comparative analyses and phylogenomic implications of *Acer* (Sapindaceae) inferred from complete chloroplast genomes. *Front Genet*. 2022;12.
43. Kim K-J, Lee H-L. Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Res*. 2004;11:247–61.
44. Zhu A, Guo W, Gupta S, Fan W, Mower JP. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytol*. 2016;209:1747–56.
45. Zhang T, Xing Y, Xu L, Bao G, Zhan Z, Yang Y, et al. Comparative analysis of the complete chloroplast genome sequences of six species of *Pulsatilla* Miller, Ranunculaceae. *Chin Med*. 2019;14:53.
46. Niu Y, Su T, Wu C, Deng J, Yang F. Complete chloroplast genome sequences of the medicinal plant *Aconitum Transsectum* (Ranunculaceae): comparative analysis and phylogenetic relationships. *BMC Genomics*. 2023;24:90.
47. Qin Z, Wang Y, Wang Q, Li A, Hou F, Zhang L. Evolution analysis of simple sequence repeats in *Plant Genome*. *PLoS ONE*. 2015;10:e0144108.
48. Ping J, Feng P, Li J, Zhang R, Su Y, Wang T. Molecular evolution and SSRs analysis based on the chloroplast genome of *Callitropsis Funebris*. *Ecol Evol*. 2021;11:4786–4802.
49. Yamamoto T. DNA markers and molecular breeding in Pear and other Rosaceae Fruit Trees. *Hortic J*. 2021;90:1–13.
50. Xia Q, Zhang H, Lv D, El-Kassaby YA, Li W. Insights into phylogenetic relationships in *Pinus* inferred from a comparative analysis of complete chloroplast genomes. *BMC Genomics*. 2023;24:346.
51. Huang X, Coulibaly D, Tan W, Ni Z, Shi T, Li H, et al. The analysis of genetic structure and characteristics of the chloroplast genome in different Japanese apricot germplasm populations. *BMC Plant Biol*. 2022;22:354.
52. Olmstead RG, Palmer JD. Chloroplast DNA systematics: a review of methods and data analysis. *Am J Bot*. 1994;81:1205–1224.
53. CHESTER M, COWAN RS, FAY MF. Parentage of endemic *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plastid DNA. *Bot J Linn Soc*. 2007;154:291–304.
54. Powell W, Morgante M, McDevitt R, Vendramin GG, Rafalski JA. Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proc Natl Acad Sci*. 1995;92:7759–7763.
55. Bock DG, Andrew RL, Rieseberg LH. On the adaptive value of cytoplasmic genomes in plants. *Mol Ecol*. 2014;23:4899–4911.

56. Huang R, Xie X, Chen A, Li F, Tian E, Chao Z. The chloroplast genomes of four *Bupleurum* (Apiaceae) species endemic to Southwestern China, a diversity center of the genus, as well as their evolutionary implications and phylogenetic inferences. *BMC Genomics*. 2021;22:714.
57. Tao L, Duan H, Tao K, Luo Y, Li Q, Li L. Complete chloroplast genome structural characterization of two *Phalaenopsis* (Orchidaceae) species and comparative analysis with their alliance. *BMC Genomics*. 2023;24:359.
58. Dong W, Liu J, Yu J, Wang L, Zhou S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS ONE*. 2012;7:e35071.
59. Park S, Jansen RK, Park S. Complete plastome sequence of *Thalictrum coreanum* (Ranunculaceae) and transfer of the rpl32 gene to the nucleus in the ancestor of the subfamily Thalictrioideae. *BMC Plant Biol*. 2015;15:40.
60. Bi D, Han S, Zhou J, Zhao M, Zhang S, Kan X. Codon usage analyses reveal the evolutionary patterns among Plastid genes of Saxifragales at a larger-sampling scale. *Genes*. 2023;14:694.
61. Yang X, Luo X, Cai X. Analysis of codon usage pattern in *Taenia saginata* based on a transcriptome dataset. *Parasites Vectors*. 2014;7:527.
62. Chakraborty S, Yengkhom S, Uddin A. Analysis of codon usage bias of chloroplast genes in *Oryza* species. *Planta*. 2020;252:67.
63. Gharib WH, Robinson-Rechavi M. The branch-site test of positive selection is surprisingly robust but lacks power under synonymous substitution saturation and variation in GC. *Mol Biol Evol*. 2013;30(7):1675–1686.
64. Fletcher W, Yang Z. The effect of insertions, deletions, and alignment errors on the branch-site test of positive selection. *Mol Biol Evol*. 2010;27(10):2257–2267.
65. Kurita M. Chromosome studies in Ranunculaceae XI. Karyotypes of seven genera. *Mem Ehime Univ*. 1958;3:13–322.
66. Al-Kelidar RK, Richards AJ. Chromosomal indications of evolutionary trends in the Genus *Delphinium* L. *Cytologia*. 1981;46:623–633.
67. Volkova SA, Ulanova KP. Chromosome numbers in some species of Nelumbonaceae and Ranunculaceae families from the Far East of the USSR. *Botanicheskii Zhurnal*. 1986;71.
68. Jian-Quan LIU, Ting-Nong H. Karyotypes of Seven Taxa of *Delphinium* from Southern Qinghai. *Plant Divers*. 1999;21:1.
69. Yuan Q, Yang Q. Low incidence of polyploids and high uniformity of karyotypes displayed by *Delphinium* (Ranunculaceae) in the Hengduan Mountains region of south-west China. *Bot J Linn Soc*. 2008;158:172–188.
70. Bobrov AA, Zalewska-Galosz J, Jopek M, Movergoz EA. *Ranunculus schmalhauseni* (section *Batrachium*, Ranunculaceae), a neglected water crowfoot endemic to Fennoscandia—a case of rapid hybrid speciation in postglacial environment of North Europe. *Phytotaxa*. 2015;233:101–138.
71. Hong Y, Luo Y, Gao Q, Ren C, Yuan Q, Yang Q-E. Phylogeny and reclassification of *Aconitum* subgenus *Lycocotnum* (Ranunculaceae). *PLOS ONE*. 2017;12:e0171038.
72. Borodina-Grabovskaya AE, Delphinium L. In: Grubov VI, editor. *Plants of Central Asia: Plant collections from China and Mongolia*, Enfield (NH). Jersey, Plymouth: Science; 2007. pp. 36–59.
73. Aboul-Maaty NA-F, Oraby HA-S. Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. *Bull Natl Res Cent*. 2019;43:25.
74. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455–477.
75. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–359.
76. Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 2020;21:241.
77. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*. 2015;31:3350–3352.
78. Qu X-J, Moore MJ, Li D-Z, Yi T-S. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods*. 2019;15:50.
79. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–1649.
80. Chan PP, Lin BY, Mak AJ, Lowe TM. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res*. 2021;49:9077–9096.
81. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res*. 2004;32(1):11–16.
82. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res*. 2019;47:W59–W64.
83. Katoh K, Standley DM. MAFFT multiple sequence alignment Software Version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–780.
84. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. Mol Biol Evol. 2017;34:3299–302. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets.
85. Peden J. Analysis of Codon Usage. The degree of Doctor of Philosophy. the University of Nottingham; 1999.
86. Sharp PM, Tuohy TMF, Mosurski KR. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res*. 1986;14:5125–5143.
87. Frank W. The effective number of codons' used in a gene. *Gene*. 1990;87:23–29.
88. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res*. 2001;29:4633–4642.
89. 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;37:1530–1534.
90. Gao F, Chen C, Arab DA, Du Z, He Y, Ho SYW. EasyCodeML: a visual tool for analysis of selection using CodeML. *Ecol Evol*. 2019;9:3891–3898.
91. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007;24:1586–1591.
92. Hurst L. The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet*. 2002;18:486.
93. Anisimova M, Bielawski JP, Yang Z. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol Biol Evol*. 2001;18:1585–1592.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.