Journal of Advanced Research 36 (2022) 175–185

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Plastid genome data provide new insights into the phylogeny and evolution of the genus *Epimedium*

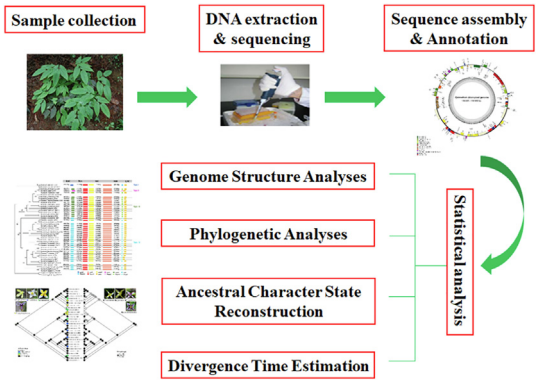
Mengyue Guo [a,b,1], Xiaohui Pang [a,b,*,1], Yanqin Xu [c], Wenjun Jiang [a,b], Baosheng Liao [d], Jingsheng Yu [a,b], Jiang Xu [d], Jingyuan Song [a,b], Shilin Chen [d,*]

[a] Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China
[b] Engineering Research Center of Chinese Medicine Resource, Ministry of Education, Beijing 100193, China
[c] College of Pharmacy, Jiangxi University of Traditional Chinese Medicine, Nanchang 330004, China
[d] Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

HIGHLIGHTS

- The phylogeny and evolution of *Epimedium* based on 45 plastomes were first studied systematically.
- Four types of plastomes were recognized according to the variation of inverted repeat boundary.
- A strong support for the sister relationship of sect. *Macroceras* and sect. *Diphyllon* was revealed.
- The diversification of *Epimedium* was estimated in the early Pleistocene (~2.11 Ma).
- Long spur (large-flowered group) constituted the plesiomorphic state of *Epimedium*.
- Disharmony existed between molecular phylogeny and traditional classification of sect. *Diphyllon*.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 14 August 2020
Revised 14 May 2021
Accepted 26 June 2021
Available online 30 June 2021

Keywords:
Epimedium
Plastid genome
Phylogeny
Divergence times
Character evolution

ABSTRACT

Introduction: *Epimedium* L., the largest herbaceous genus of Berberidaceae, is one of the most taxonomically difficult representatives. The classification and phylogenetic relationships within *Epimedium* are controversial and unresolved.
Objectives: For the first time, we systematically studied the phylogeny and evolution of *Epimedium* based on plastid genome (plastome) data for better understanding this enigmatic genus.
Methods: We explored the molecular phylogeny, assessed the infrageneric classification, estimated the divergence times, and inferred the ancestral states for flower traits of *Epimedium* based on 45 plastomes from 32 species.
Results: The *Epimedium* plastome length ranged from 156,635 bp to 159,956 bp. Four types of plastome organization with different inverted repeat boundary changes were identified. Phylogenetic analysis revealed a strong support for the sister relationship of sect. *Macroceras* and sect. *Diphyllon* but did not provide a distinct route for petal evolution in *sect. Diphyllon*. Disharmony between phylogenetic relationships and traditional classification of sect. *Diphyllon* was observed. Results from divergence time analysis

Peer review under responsibility of Cairo University.
* Corresponding authors at: Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China (X. Pang); Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China (S. Chen).
E-mail addresses: xhpang@implad.ac.cn (X. Pang), slchen@icmm.ac.cn (S. Chen).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jare.2021.06.020

showed that *Epimedium* diverged in the early Pleistocene (~2.11 Ma, 95% HPD = 1.88–2.35 Ma). Ancestral character state reconstructions indicated transitions from long spur (large-flowered group) to other petal types (small-flowered group) in *Epimedium*.

Conclusion: These findings provide new insights into the relationships among *Epimedium* species and pave the way for better elucidation of the classification and evolution of this genus.

Introduction

Epimedium L., a medicinally important and highly speciose genus of family Berberidaceae, comprises approximately 60 species, which are distributed around the world [1–3]. Dispersing from Japan to Algeria, *Epimedium* is unevenly distributed in eastern Asia and the Mediterranean region [1,2]. The species diversity of *Epimedium* is highest in China, where approximately 50 species have been reported [4,5]. Eastern China, which serves as a museum for herbaceous genera, is the main distribution region of *Epimedium* species [6]. Many *Epimedium* species have been used to treat sexual dysfunction, osteoporosis, and cardiovascular diseases in China, Korea, and Japan for a long time [7]. Considering their great medicinal value, researchers have studied *Epimedium* species in depth.

A long period of time has passed from the first record of *Epimedium* species [8] to the establishment of the most comprehensive classification system for this genus [1]. According to Stearn's monograph [1], *Epimedium* is divided into two subgenera, four sections, and four series based on leaf and flower morphology, geographical distribution, and C-banding of chromosomes. Subg. *Rhizophyllum* comprises two species in Algeria and Caucasia. Subg. *Epimedium* consists of four sections, namely, *Epimedium* (two species; Alps and Balkan areas and Caucasia), *Polyphyllon* (one species; western Himalaya), *Macroceras* (six species; Japan, Korea, northeastern China, and Far Eastern Russia), and *Diphyllon* (43 species; central-southeastern China). Sect. *Diphyllon* is subdivided into four series, namely, *Campanulatae*, *Davidianae*, *Dolichocerae*, and *Brachycerae*, based on flower morphology, particularly petal characteristics [1]. Both of ser. *Campanulatae* and ser. *Brachycerae* have small flowers, and they were differentiated by the flat petals in the former and the very short petals (saccate or short spurs) in the latter. Ser. *Davidianae* and ser. *Dolichocerae* both have large flowers and petals with long spurs, and the two are distinguished by the occurrence and disappearance of basal laminae [1]. Subsequent phylogenetic studies based on karyotypes and molecular markers could not completely confirm the widely accepted classification system of Stearn's [6,9–13]. Subg. *Rhizophyllum* and four sections of subg. *Epimedium* were consistently supported as five distinct clades [6,9,12,13]. However, the two subgenera have been proven to be paraphyletic by Zhang et al. [6] and De Smet et al. [13]. The relationships of the four sections are poorly supported, except for the sect. *Epimedium*, which is the sister of sect. *Macroceras* [6,9,13]. The subdivision of sect. *Diphyllon* could not be well supported as species from the same series did not cluster into one clade [6,9,12,13]. Meanwhile, the petal evolution routes of *Epimedium* have not been resolved [9,12,13], and further exploration on the evolution and origin of this genus is needed.

Some taxonomic questions on *Epimedium* species are up for debate because of the morphological variation within and among some species. Many *Epimedium* species described at the early stages were only based on a single locality, and a limited number of specimens were used. Shortage of adequate investigation and high-quality specimens hinders the accurate and detailed description of one species [4,14–16], which might impede the assignment of specimens to the recognized taxa, thereby leading to the unnecessary publication of new species and complicating the taxonomic relationships among morphologically similar species [13]. On the one hand, some new species that have been published are synonymous with existing species. *Epimedium chlorandrum*, *E. rhizomatosum*, *E. brachyrrhizum*, *E. dewuense*, and *E. sagittatum* var. *oblongifoliolatum* were respectively treated as the synonyms of *E. acuminatum*, *E. membranaceum*, *E. leptorrhizum*, *E. dolichostemon*, and *E. borealiguizhouense* [3]. On the other hand, the existence of species complex, which is referred as a cluster of closely related species, complicates the classification issues of *Epimedium*. For instance, *E. sagittatum* complex, which consists of five species and three varieties, is the most controversially defined taxa in this genus [12]. Many species are differentiated based on small variations of leaf and flower characters, the boundaries between some taxa are blurred. Therefore, effective methods to resolve phylogenetic relationships and assess previous classification of *Epimedium* species are urgently needed.

The plastid genomes (plastomes), generally ranging from 120 to 160 kb, are highly conserved in terms of structure, size, and gene content in angiosperms [17,18]. The plastomes can increase phylogenetic resolution at low taxonomic levels in plant phylogenetic analyses [19], the results of which are used to resolve phylogenetic relationships of plant groups [20–22]. For example, Niu et al. [20] found that *Triplostegia* is composed of four main clades that largely correlate with geography. The research of Zhong et al. [22] supported the monophyly of section *Obconicolisteri* and inferred multiple independent transitions from distyly to homostyly within this section. Improved sequencing techniques in combination with the widespread interest in plastomes have led to the considerable increase of plastome publishing, with more than 3000 available in GenBank [23]. The first plastome of *Epimedium* species (*E. koreanum*) was reported by Lee et al. [24]. Later, other four cp genomes from *E. acuminatum*, *E. dolichostemon*, *E. lishihchenii*, and *E. pseudowushanense* were published by Zhang et al. [6] and Sun et al. [25]. Subsequently, our group reported nine new plastomes from *E. wushanense* and its closely related species [26]. These studies initially explored the structural patterns of *Epimedium* plastomes and the phylogenetic relationships among some species. However, the existing data are not sufficient to comprehensively illustrate the intricate phylogenetic relationships within the entire genus. Thus, more evidence is indispensable for further investigation.

For the first time, we used 45 plastid genomes from 32 *Epimedium* species to reconstruct the phylogeny. This study aimed to do the following: (1) examine infrageneric classification; (2) estimate divergence times; and (3) trace the evolution of the floral phenotypes in Chinese *Epimedium*. The results obtained in this study can improve our understanding of the classification, phylogeny, and evolution of this important and enigmatic genus.

Materials and methods

Taxon sampling, DNA Extraction, and sequencing

We collected 45 individuals from 32 *Epimedium* species. Among them, 32 plant materials were newly sampled from Jilin, Hubei,

Hunan, Anhui, Shaanxi, Guizhou, Chongqing, and Sichuan Provinces of China (Table S1). These samples were all identified by taxonomists Prof. Yanqin Xu at Jiangxi University of Traditional Chinese Medicine and Prof. Shunzhi He at Guiyang College of Traditional Chinese Medicine. Thirteen plastomes previously published [6,25,26] were also included. Detailed information is provided in Table S1, which is ordered according to the classification of Stearn [1]. In addition, two species of *Vancouveria* were included as outgroups, since this genus is most closely related to *Epimedium* [6] (Table S2).

The fresh leaves were frozen at −20 °C before DNA extraction. Total genomic DNA was extracted and purified using the Plant Genomic DNA Rapid Extraction kit (Bioteke Corporation, Beijing, China) in accordance with the manufacturer's instructions. The quantified DNA was used to construct shotgun libraries with average insert sizes of 500 bp and sequenced using Illumina Hiseq X platform (Illumina, San Diego, CA, USA) in accordance with the manufacturer's manual. Raw data from each sample were produced with 150 bp pair-end read lengths.

Sequence assembly and annotation

Low-quality reads from raw data were filtered using the software Skewer v. 0.2.2 [27]. BLAST search was performed against the available *Epimedium* plastomes to select the contigs of plastid origin. ABySS 2.0.0 program [28] was used for initial assembly using a kmer size of 127. The assembled contigs were compared with the reference *Epimedium* genome using MUMmer v. 3.0 [29] to achieve the correct order of the contigs. Gaps in the assemblies were filled by Sanger sequencing with specific primers designed for PCR. To ensure assembly accuracy, the four junctions between the inverted repeats (IRs) and single-copy (SC) regions were verified through PCR amplification and Sanger sequencing with specific primers. All the primers used are listed in Table S3.

Plastome annotation was performed using CPGAVAS2 [24]. The plastome maps were generated using the Organellar-Genome DRAW v. 1.3.1 [30] with default settings and were checked manually. The newly gained plastome sequences were deposited in GenBank under accession numbers MT560392-MT560423 (Table S1).

Genome structure analyses

The GC content was analyzed using the software Mega 6.0 [31]. Simple sequence repeats (SSRs) were identified using the MISA Perl script (http://pgrc.ipk-gatersleben.de/misa/). Microsatellites were detected with thresholds of 10 repeat units for mono-, six repeat units for di-, four repeat units for tri- and tetra-, and three repeat units for penta- and hexanucleotide SSRs. The mVISTA program was used to compare the 45 *Epimedium* genomes in Shuffle-LAGAN mode [32]. DNA polymorphism analysis was performed using DnaSP (DNA Sequence Polymorphism) v6 [33] to calculate the nucleotide diversity (Pi) and to detect highly variable sites among *Epimedium* plastomes with a step size of 200 bp and window length of 800 bp. The 45 whole plastomes were also aligned to assess possible rearrangements in GENEIOUS via the MAUVE plugin [34].

Phylogenetic analyses and ancestral character state reconstruction

Phylogenetic relationships of *Epimedium* were inferred using maximum likelihood (ML) and Bayesian inference (BI) methods. Three data sets were generated for phylogenetic analyses, as follows: (1) 45 whole plastomes; (2) a concatenated data set consisting of coding sequence (CDS) of the 77 protein-coding genes (excluding *rpl32* gene); and (3) the non-coding regions. Two species belonging to the genus *Vancouveria* were chosen as outgroups to root the tree (Table S2). Sequences were aligned using MAFFT v. 7.215 [35] and trimmed using trimAl v. 1.4 with option -automated1 [36]. The ML analyses were performed using IQ-TREE v. 1.6.12 [37] with 10,000 bootstrap replicates. The best fitting model was selected by ModelFinder [38] implemented in IQ-TREE. The BI analysis was performed using MrBayes v. 3.2.2 [39] under a GTR + I + G model selected by Akaike information criterion (AIC) in MrModeltest 2.3 [40]. The analyses were performed with four parallel Markov Chain Monte Carlo (MCMC) runs for two million generations, sampling every 100 generations and with a 25% burn-in. Trees were visualized in FigTree 1.4.3 [41].

Taxonomically important morphological characters, such as petal shape and flower diameter, were selected to analyze the morphological evolution of the *Epimedium* genus. Morphological information was obtained from the description in the Flora of China [2] and taxonomic literature [16,42–44]. States for the two characters were defined based on the classification system of Stearn [1] as follows: petal shape (scheme 1): (0) flat, (1) saccate or short spur, (2) long spur without basal laminae, (3) long spur with basal laminae; flower diameter: (0) small flower (shorter than inner sepals, <1 cm in diameter), (1) large flower (longer than inner sepals, >1 cm in diameter). We also defined the character states for petal shape (scheme 2) according to Ying's [45] classification, as follows: (0) flat, (1) slightly saccate, (2) saccate, (3) short spur, and (4) long spur. The number of individuals used for ancestral character reconstruction was reduced to 32, consistent with the number of species. The ML tree was constructed using IQ-TREE v. 1.6.12 [37] with 10,000 bootstrap replicates. The ancestral character state reconstruction analysis was performed over the ML tree obtained with the Mk1 model in Mesquite v. 3.6 [46]. The ML and parsimony criterion were used for character optimization.

Divergence time estimation

BEAST v. 2.5.1 was used to estimate the divergence times [47]. The CDS of the 77 protein-coding genes from 45 individuals of *Epimedium* and 28 species belonging to Ranunculales were involved in the analysis (Table S2). The nucleotide substitution model GTR + G was selected. A strict clock analysis with a Yule process speciation model was specified. We selected two fossils as internal calibration points: (1) The split between *Ranzania* and *Mahonia-Berberis* was constrained to be 33.8 Ma based on the fossils of *Mahonia* [48–50]. (2) The split between *Eranthis* and *Actaea* was constrained to be 55.8 Ma based upon the fossil of *Paleoactaea nagelii* [50,51]. Log-normal distributions with 1.0 standard deviation were used for the two fossil points. The crown group of Ranunculales was constrained to a minimum age of 112.0 Ma [50,52]. The root of the tree was calibrated using a uniform distribution between 112.0 and 124.0 Ma [50]. MCMC chains were run for 10 million generations, sampling every 1,000 generations. Tracer v. 1.6 [53] was used to assess the convergence according to the criterion of effective sample sizes (ESS) > 200. A maximum clade credibility (MCC) tree with median values was produced after a burn-in of 25% using TreeAnnotator v. 2.5.1. The chronogram was visualized in FigTree 1.4.3 with 95% highest posterior density (95% HPD) for each node [41].

Results

Structure features of Epimedium plastomes

The length of assembled *Epimedium* plastomes ranged from 156,635 bp (*E. myrianthum*, GZSS) to 159,956 bp (*E. davidii*, SCFZ) (Fig. 1). All the plastomes showed a typical quadripartite structure comprising a large single-copy (LSC) region (85,862–89,643 bp)

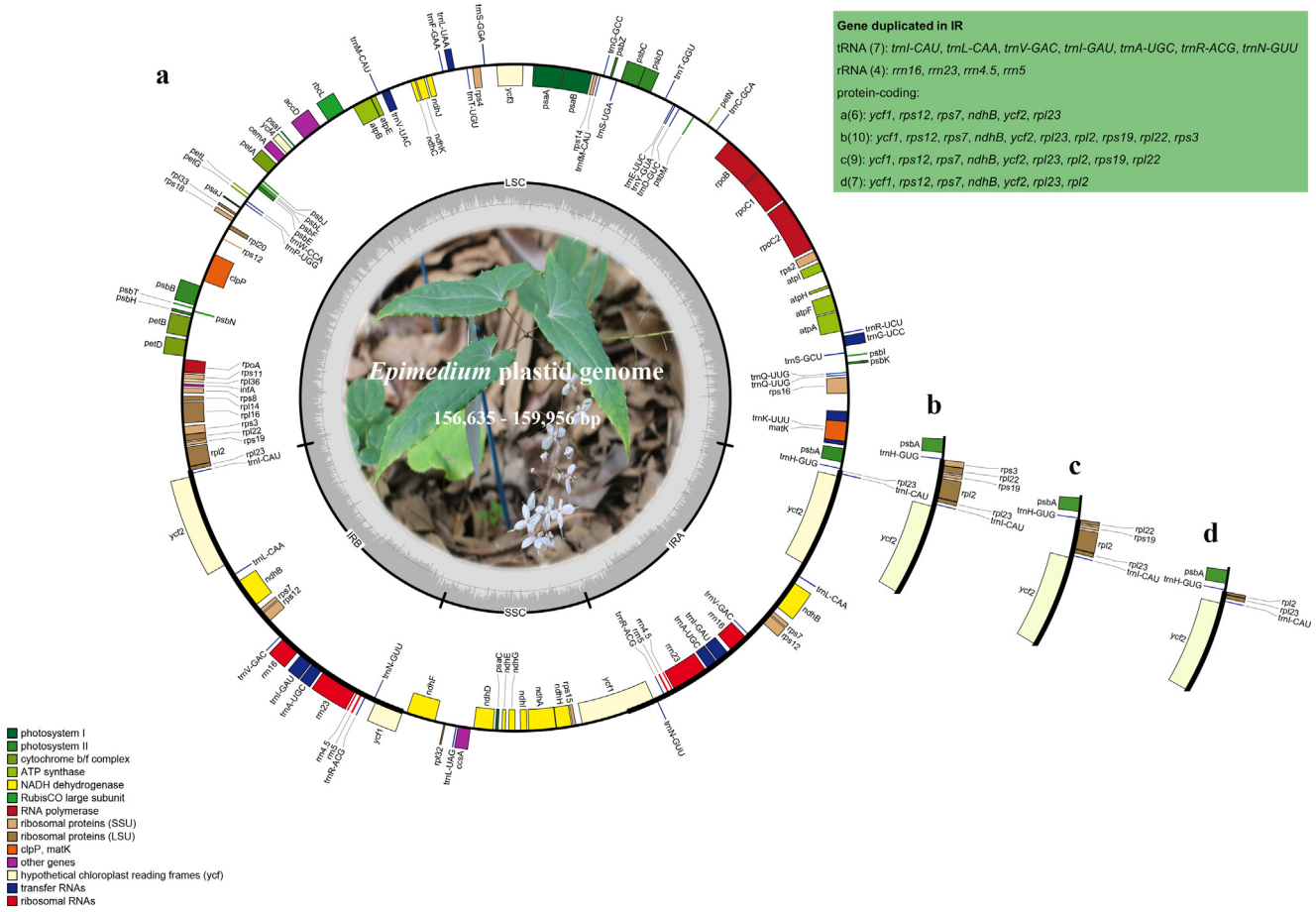


Fig. 1. Gene map of the *Epimedium* plastomes. (a–d) Type I-IV plastomes represented by that of *E. koreanum*, *E. davidii* (SCFZ), *E. brevicornu*, and *E. wushanense*. Genes drawn inside the circle are transcribed clockwise, whereas those outside the circle are transcribed counterclockwise.

and a small single-copy (SSC) region (16,103–17,216 bp) separated by two inverted repeat (IR) regions (25,218–28,506 bp) (Fig. 1; Table 1). The average GC content is ~38.8% (Table1). The total number of unique gene annotated is 112, comprising 78 protein-coding, 30 tRNA, and 4 rRNA genes (Fig. 1).

The number of SSRs identified in 40 *Epimedium* plastomes ranged from 69 (*E. pubescens*, SCYA) to 84 (*E. koreanum* and *E. epsteinii*) (Table S4). Six kinds of repeat patterns were screened. Among the mononucleotide repeats, A/T was abundant (81.8–89.9%), whereas C/G was rare (2.4–4.2%). Dinucleotides (1.2–3.8%) and trinucleotides (3.8–8.4%) were identified in all samples. Pentanucleotides (0–1.4%) occurred in 25 samples, whereas tetranucleotides (0–1.4%) and hexanucleotides (0–3.9%) were found in only five and eight samples, respectively (Table S4). SSRs were primarily distributed in LSC and SSC region. Most SSRs were identified in the intergenic regions, followed by introns and exons. These SSRs could be used to develop potential molecular markers for species differentiation and population genetics in future research.

High conservation of plastomes in *Epimedium* was revealed using mVISTA (Fig. S1). The alignment showed great similarity (99.0–99.9%) within the genus with *E. wushanense* (sample HBXS_1) as the reference. The non-coding regions were more divergent than the coding regions. Four intergenic spacer regions, namely, *psbC-trnS-psbZ*, *accD-psaI*, *ndhF-rpl32-trnL*, and *ndhD-psaC* were observed, and these had higher Pi values than the other regions (Pi > 0.005) (Fig. S2). These divergent regions contain abundant variation information and can be used to develop molecular markers as potential DNA barcodes for *Epimedium* species authentication. Moreover, the whole genome alignment analysis results showed that no genomic rearrangements were detected (Fig. S3).

Phylogenetic analyses

Six phylogenetic trees based on ML and BI analyses of the 77 protein-coding genes, the non-coding regions, and whole plastomes were constructed. The protein-coding data set had an aligned length of 68,730 nucleotide sites, of which 888 nucleotide sites (1.29%) were parsimony informative among *Epimedium* species. The non-coding data set comprised 68,163 nucleotide sites, 2,051 (3.01%) of which were parsimony informative. The whole plastome data set had an aligned length of 159,269 nucleotide sites, among which 2,927 (1.84%) were parsimony informative characters. The topologies obtained using ML and BI methods were nearly identical based on each data set (Figs. S4–S6). Topological conflicts between the non-coding and the other two data sets were observed. Considering the protein-coding and whole plastome data sets yielded better supported trees in comparison with the non-coding data set, the phylogenetic relationships were discussed mainly based on the results obtained from the former. Apart from some weakly supported nodes (*E. epsteinii*, *E. ilicifolium*, *E. sagittatum*, *E. wushanense*, *E. glandulosopilosum*, and *E. zhushanense*), largely congruent topologies were produced by the protein-coding and whole plastome data sets (Figs. S4–S5). Two strongly supported clades were recognizable, and they corresponded to sect. *Macroceras* (Clade A) and sect. *Diphyllon* (Clade B), as classified

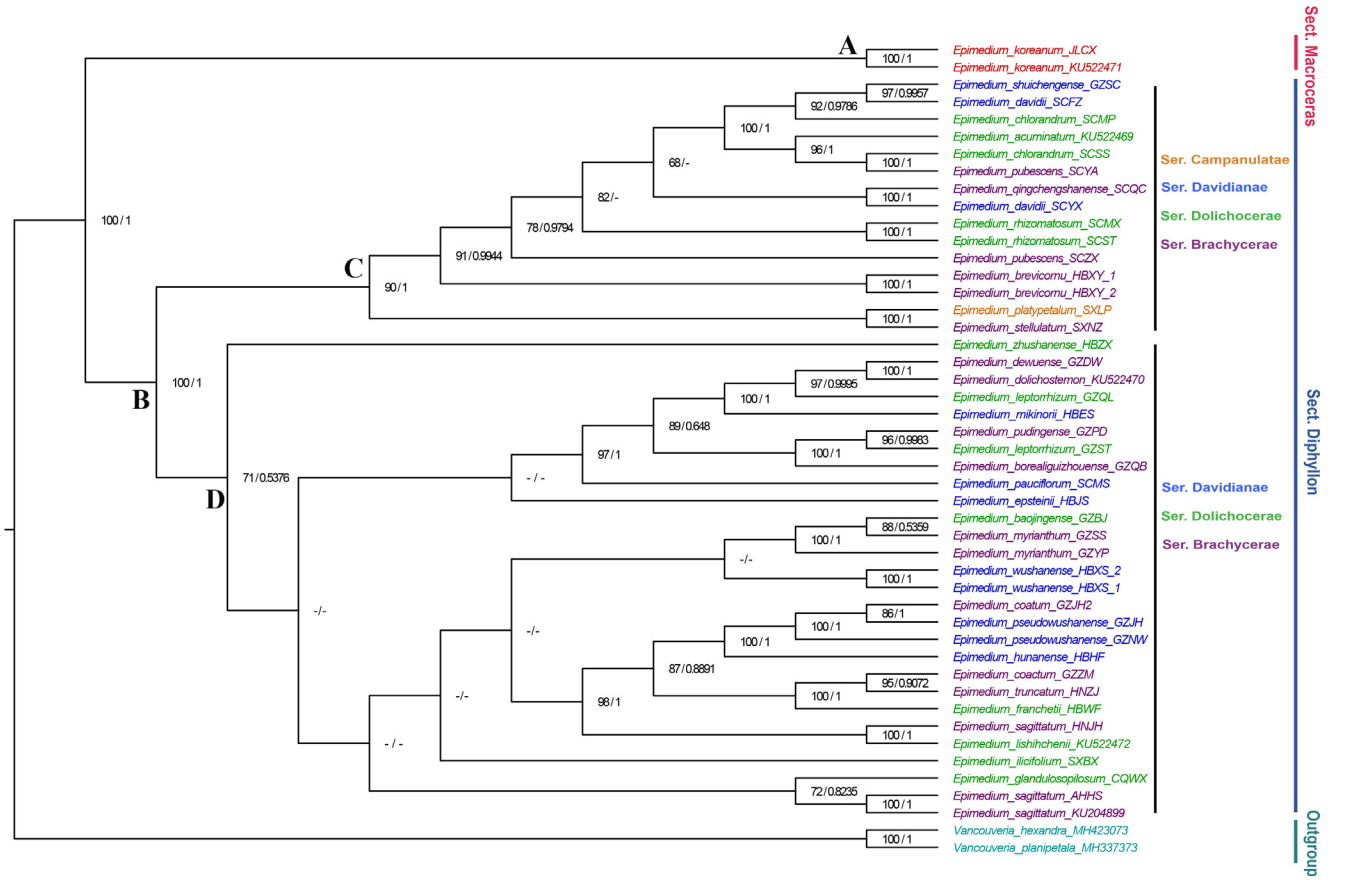


Fig. 2. Phylogenetic tree reconstructed based on the CDS of the 77 protein-coding genes using maximum likelihood (ML) and Bayesian interference (BI) methods. Numbers at the nodes represent ML bootstrap (BS) and BI posterior probability (PP) values. BS or PP values lower than 50% or 0.5 were indicated by hyphens.

by Stearn [1]. The sect. *Diphyllon* further diversified into two clades, and species belonging to different series [1] were nested together in each clade. Clade C included all the four series, namely, ser. *Campanulatae*, ser. *Davidianae*, ser. *Dolichocerae*, and ser. *Brachycerae*. Clade D comprised ser. *Davidianae*, ser. *Dolichocerae*, and ser. *Brachycerae*. However, the support for further division of Clade D into two subclades was low. For species with multiple populations, the monophyly of four species (*E. koreanum*, *E. wushanense*, *E. brevicornu*, and *E. rhizomatosum*) was well supported. However, the different populations of *E. davidii*, *E. chlorandrum*, *E. pubescens*, *E. leptorrhizum*, *E. coatum*, and *E. sagittatum* were dispersedly nested.

Expansion and contraction of the IR region

High synteny of genes arrangement in *Epimedium* plastomes was observed apart from the extent of the IR regions. The differences in plastome size of *Epimedium* are primarily attributed to the extent of the IR region (Fig. 3). All *Epimedium* plastomes have the SSC/IRB boundary within the *ycf1* gene and the SSC/IRA border between the pseudogene (ψ) *ycf1* and gene *ndhF*. Four types of plastomes were characterized by IR/LSC boundary variation (Figs. 1 and 3). Type I plastomes were identified in Clade A (sect. *Macroceras*), and these were characterized by the IRA/LSC boundary within the coding region of *rpl23* and the IRB/LSC junction between ψ*rpl23* and *trnH-GUG*. The remaining three types of plastomes were identified in Clade B (sect. *Diphyllon*), with IR region expanding by approximately 0.5–3.3 k to the LSC region. Types II and III corresponded to Clade C, whereas type IV occurred in Clade D. The IRA/LSC boundaries of Types II, III, and IV expanded to the coding region of *rps3*, *rpl22*, and *rpl2*, respectively. The IRB/LSC borders for these three types of plastomes were correspondingly located between ψ*rps3* and *trnH-GUG*, ψ*rpl22* and *trnH-GUG*, and ψ*rpl2* and *trnH-GUG*.

In each type of plastome, the same number of duplicated rRNA and tRNA genes (four and seven) and different number of duplicated protein-coding genes were found (Fig. 1). Type I and II were respectively found only in *E. koreanum* and *E. davidii* (sample SCFZ), and these presented the least (17) and most (21) duplicate genes in the IRs (Fig. 1a, b; Table 1). Type III was present in 10 species with 20 duplicated genes (Fig. 1c; Table 1). Type IV was found in 21 species with 18 duplicated genes in the IRs (Fig. 1d; Table 1).

Morphological character evolution

The reconstruction of the petal shape (scheme 2) in *Epimedium* indicated that the ancestral state is long spur (Fig. 4a). Short spur and saccate evolved at least three and four times in Clade B, respectively. The slightly saccate evolved at least four times in Clade D. Lastly, *E. platypetalum* is flat. Our data suggested that the other four petal types evolved from the long spur, and no reversals back to long spur occurred. Our analysis revealed large flower (longer than inner sepals, >1 cm in diameter) as the ancestral character state for *Epimedium* (Fig. 4b). However, ancestral state reconstruction for petal shape (scheme 1) showed that this character exhibited rampant diversity among the major clades and its ancestral state was equivocal (Fig. S7).

Table 1
Characteristics of the *Epimedium* plastomes generated in this study.

Latin name	Voucher No.	LSC length/bp	SSC length/bp	IR length/bp	Genome size/bp	GC content/%	Genome type	Gene duplicated in IR (protein-coding/tRNA/rRNA)
E. koreanum	JLCX	89,643	17,216	25,218	157,295	38.71	I	6/7/4
E. davidii	SCFZ	85,862	17,082	28,506	159,956	38.82	II	10/7/4
E. davidii	SCYX	86,619	17,064	27,716	159,115	38.8	III	9/7/4
E. brevicornu	HBXY_1	86,525	17,015	27,698	158,936	38.82	III	9/7/4
E. brevicornu	HBXY_2	86,525	17,015	27,698	158,936	38.82	III	9/7/4
E. chlorandrum	SCMP	86,626	16,237	27,694	158,251	38.89	III	9/7/4
E. chlorandrum	SCSS	86,542	17,084	27,700	159,026	38.82	III	9/7/4
E. platypetalum	SXLP	86,542	17,050	27,851	159,294	38.82	III	9/7/4
E. shuichengense	GZSC	86,627	17,070	27,759	159,215	38.8	III	9/7/4
E. stellulatum	SXNZ	86,561	17,055	27,730	159,076	38.81	III	9/7/4
E. pubescens	SCYA	86,527	17,083	27,715	159,040	38.82	III	9/7/4
E. pubescens	SCZX	86,399	17,037	27,769	158,974	38.83	III	9/7/4
E. qingchengshanense	SCQC	86,608	17,068	27,709	159,094	38.81	III	9/7/4
E. rhizomatosum	SCST	86,572	17,056	27,733	159,094	38.82	III	9/7/4
E. rhizomatosum	SCMX	86,577	17,056	27,733	159,098	38.82	III	9/7/4
E. baojingense	GZBJ	88,406	17,039	25,778	157,001	38.79	IV	7/7/4
E. borealiguizhouense	GZQB	88,602	17,039	25,820	157,281	38.78	IV	7/7/4
E. coactum	GZJH2	88,529	17,088	25,784	157,185	38.78	IV	7/7/4
E. coactum	GZZM	88,301	17,048	25,833	157,015	38.79	IV	7/7/4
E. dewuense	GZDW	88,394	17,090	25,784	157,052	38.81	IV	7/7/4
E. epsteinii	HBJS	88,542	17,058	25,790	157,180	38.8	IV	7/7/4
E. franchetii	HBWF	88,452	17,095	25,833	157,213	38.79	IV	7/7/4
E. glandulosopilosum	CQWX	88,541	17,062	25,784	157,171	38.79	IV	7/7/4
E. hunanense	HBHF	88,366	17,066	25,916	157,264	38.81	IV	7/7/4
E. ilicifolium	SXBX	88,216	17,037	25,801	156,855	38.77	IV	7/7/4
E. leptorrhizum	GZQL	88,447	17,077	25,784	157,092	38.8	IV	7/7/4
E. leptorrhizum	GZST	88,590	17,034	25,820	157,264	38.77	IV	7/7/4
E. mikinorii	HBES	88,359	17,074	25,776	156,985	38.81	IV	7/7/4
E. myrianthum	GZYP	88,376	17,024	25,777	156,954	38.8	IV	7/7/4
E. myrianthum	GZSS	88,062	16,877	25,848	156,635	38.79	IV	7/7/4
E. pauciflorum	SCMS	88,600	17,046	25,692	157,030	38.78	IV	7/7/4
E. pseudowushanense	GZJH	88,448	17,069	25,780	157,077	38.78	IV	7/7/4
E. pseudowushanense	GZNW	88,537	17,077	25,789	157,193	38.78	IV	7/7/4
E. pudingense	GZPD	88,611	17,052	25,820	157,303	38.77	IV	7/7/4
E. sagittatum	AHHS	88,437	17,054	25,780	157,051	38.81	IV	7/7/4
E. sagittatum	HNJH	88,564	16,103	26,589	157,845	38.76	IV	7/7/4
E. truncatum	HNZJ	88,316	17,061	25,833	157,043	38.8	IV	7/7/4
E. wushanense	HBXS_2	88,577	17,090	25,784	157,235	38.79	IV	7/7/4
E. wushanense	HBXS_1	88,526	17,091	25,789	157,195	38.8	IV	7/7/4
E. zhushanense	HBZX	88,377	16,986	25,930	157,223	38.81	IV	7/7/4

Molecular dating

We estimated the divergence times of the major clades within Berberidaceae based on the CDS of 77 protein-coding genes. The divergence between *Vancouveria* and *Epimedium* was estimated to occur in the late Miocene (~6.59 Ma, 95% HPD = 6.12–7.06 Ma) (Fig. 5). *Epimedium* diverged in the early Pleistocene (~2.11 Ma, 95% HPD = 1.88–2.35 Ma). Within the genus *Epimedium*, sect. *Diphyllon* probably originated at 1.14 Ma (95% HPD: 1.05–1.23 Ma), and *E. koreanum* belonging to sect. *Macroceras* arose at around 0.06 Ma (95% HPD: 0.02–0.13 Ma).

Discussion

Phylogenetic and taxonomic inference of Chinese Epimedium

Congruent with previous studies that used different molecular markers (ITS in Sun et al. [9]; ITS and *atpB-rbcL* in Zhang et al. [6]; ITS, *trnK-matK* and AFLP datasets in De Smet et al. [13]; AFLP in Zhang et al. [12]), the phylogenetic trees reconstructed based on the plastome resources strongly supported the sister relationship between sect. *Macroceras* and sect. *Diphyllon*, but did not provide resolution within sect. *Diphyllon* (Fig. 2). Here, the sect. *Diphyllon* was subdivided into two well-supported clades, and the species belonging to different series were nested together in each clade. Moreover, the species were separated with short branches in our analysis (Figs. S4–S6 and 5), supporting the hypothesis that the observed polytomy resulted from a recent radiation in the distribution area of Chinese *Epimedium* taxa [6,13]. Therefore, the common ancestor of *Epimedium* probably underwent rapid speciation without sufficient time to accumulate informative mutations [54]. Given that sect. *Diphyllon* might still be in the process of differentiation [10], intensive variation investigations are essential to document the process of species differentiation for the completion of species descriptions. The discrimination of *E. dewuense* and *E. dolichostemon* is only based on the differences of shape and indumentum of their leaflets. Extensive investigations by Zhang et al. [3] revealed that the leaflet variation between the two species was continuous, and thus, *E. dewuense* was treated as a synonym of *E. dolichostemon*. In our study, *E. dewuense* was strongly supported to be a sister to *E. dolichostemon* (BS/PP = 100%/1) and in turn supporting the previous report of Zhang et al. [3]. The clearly identified boundaries of *Epimedium* species might be the precondition to resolve the discordance and complexity of this genus.

Despite the incongruence between morphology and phylogenetic relationships, occasional evidence of geographic patterning within the genus was found. Some sister-terminals from different

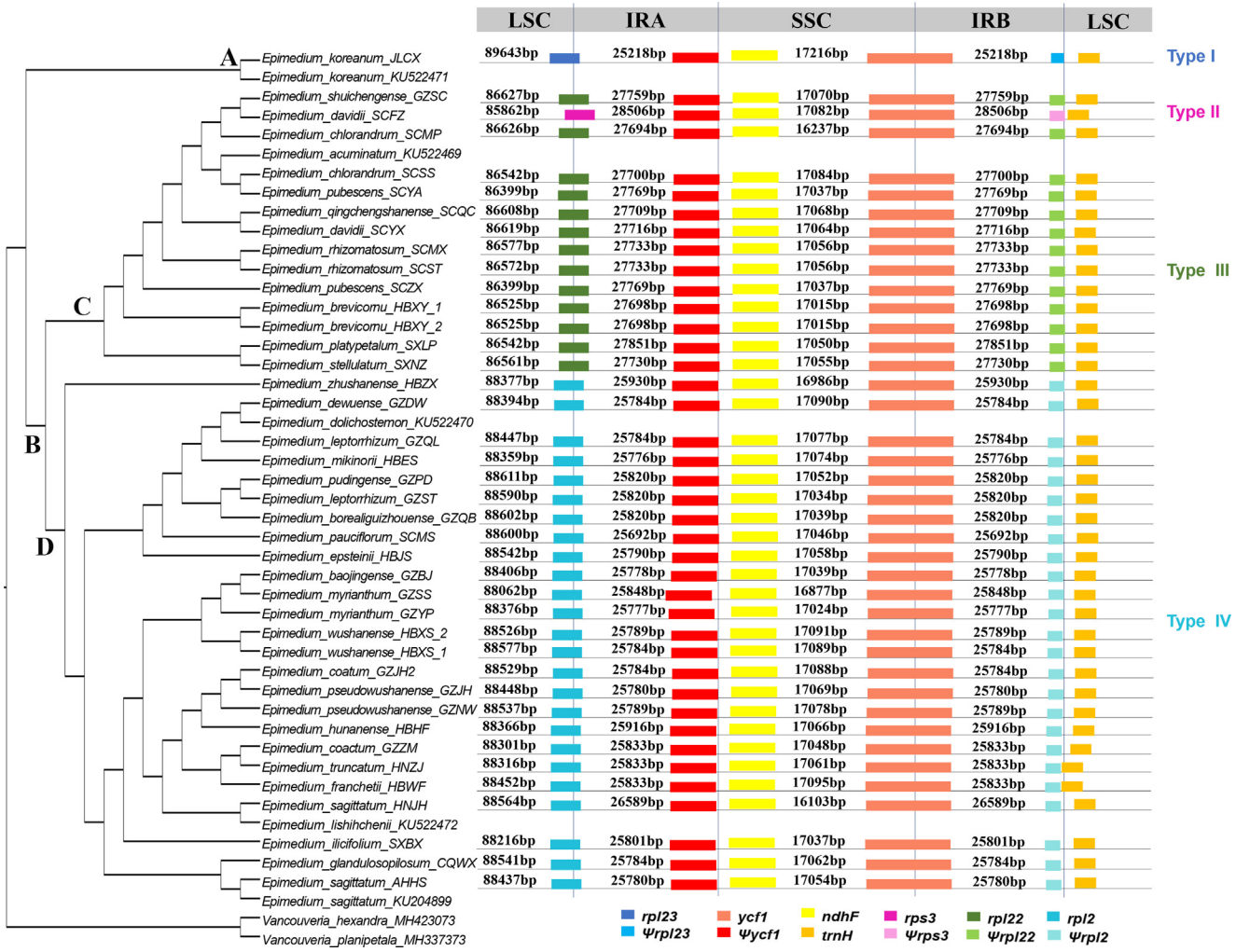


Fig. 3. Comparison of boundaries of the large single-copy, small single-copy, and inverted repeat regions among *Epimedium* plastomes.

species with distinctly morphological characteristics were collected from geographically close locations. For instance, *E. platypetalum* (ser. *Campanulatae*) and *E. stellulatum* (ser. *Brachycerae*), both collected from Nanzheng County, Shaanxi Province, formed a well-supported clade (BS/PP = 100%/1). Within Clade C, closer genetic relationship was found among different species from same geographical areas rather than among the same species from distinct areas in the individuals collected from Sichuan province (designated as SC). One population of *E. davidii* (SCFZ) and *E. chlorandrum* (SCMP), both collected from Baoxing County, Ya'an City, Sichuan Province, formed a subclade. This subclade is sister to a subclade consisting of another population of *E. chlorandrum* (SCSS from Lushan County, Ya'an City, Sichuan Province), *E. pubescens* (SCYA from Ya'an City, Sichuan Province), and *E. acuminatum* (KU522469 from Baoxing County, Ya'an city, Sichuan Province). These two subclades consisted of geographically adjacent individuals and formed a well-supported clade (BS/PP = 100%/1). Similar pattern was also observed in *E. sagittatum* and its related species, known as the *E. sagittatum* complex. The *E. sagittatum* complex is the most controversially defined taxonomy in *Epimedium* due to transitions and intersections of morphological traits among the species. Samples of four species in the abovementioned complex, namely, *E. sagittatum*, *E. coactum*, *E. myrianthum*, and *E. pudingense*, were collected in our study. We found that the species of complex were separately clustered with other species that are geographically close. The geographic pattern of shared ancestry might be attributed to gene flow between species [55], as reported for *Eucalyptus* [56] and *Silene* [57]. The interspecies similarity of *Epimedium* was close to or even higher than intraspecies similarity [15]. Frequent gene flow among sympatric populations due to the dispersal of pollen or seeds might be one of the reasons for the low interspecific differentiation of *Epimedium* species [58,59].

To conclude, the phylogenetic relationship of *Epimedium* was difficult to explain. In view of the widespread distribution of many *Epimedium* species and the overlapping distribution areas among species, further phylogeographic studies based on more populations are needed.

IR extension in Epimedium species

The contraction and expansion at the IR boundary are common evolutionary events and are the main reasons for the length variation of plastomes [60]. High conservation of gene content and structure in *Epimedium* plastomes was observed in our study (Figs. S1 and S3). However, structural changes in the IR/LSC boundary led to the variation of genome size (156,635–159,956 bp) and deletion of single copy gene. The *Epimedium* plastomes were divided into four types according to gene organization. Type I plastomes were identified in Clade A (sect. *Macroceras*), which presented the least (17) duplicate genes in the IRs (Figs. 1a, 3). The remaining three types of plastomes were found in sect. *Diphyllon*. Types II and III occurred in Clade C, with the IR region expanding

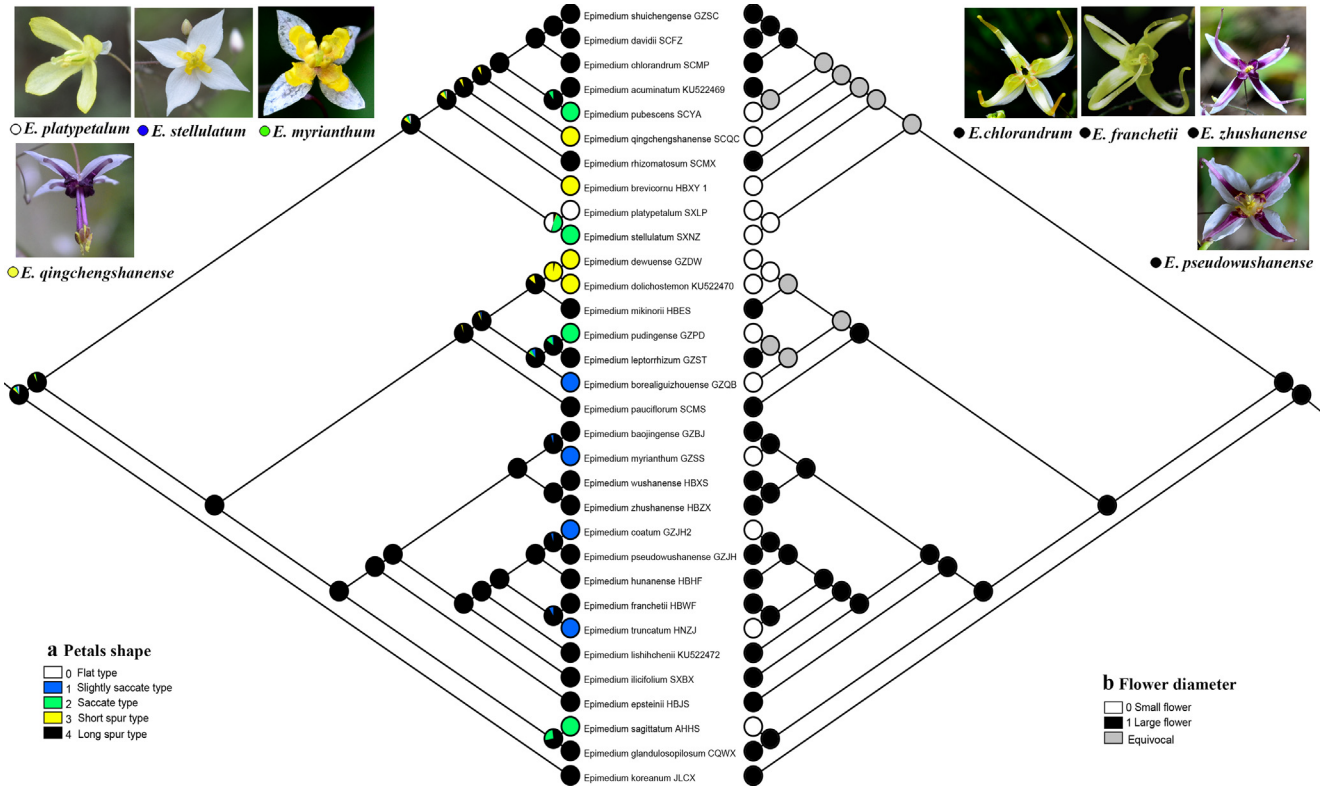


Fig. 4. Ancestral state reconstructions of petal shape (a) and flower diameter (b) in *Epimedium*. The likelihood of occurrence of each state at each node is indicated by circles made of colored wedges of different size. The images on the upper left and upper right show the characteristics of the small-flowered (flat, slightly saccate, saccate, and short spur) and large-flowered (long spur) groups, respectively.

by approximately 2.5–3.3 k to the LSC region and containing up to 21 duplicate genes (Figs. 1b–c, 3). Type IV was identified in Clade D, with IR region extending by about 0.5–1.4 k to the LSC region and presenting 18 duplicate genes (Figs. 1d, 3).

Dating the origin of diversification of Chinese Epimedium

Consistent with the previous findings [61,62], the sister relationship between *Epimedium* and *Vancouveria* was confirmed. The estimated age for their split dated back to the late Miocene (~6.59 Ma, 95% HPD = 6.12–7.06 Ma), as shown in Fig. 5. The uplifts of Qinghai-Tibet Plateau (QTP) in association with the Quaternary climatic oscillations have been discussed in shaping geographic genetic structure and triggering speciation and diversifications [63]. The diversification of *Epimedium* in China was estimated in the early Pleistocene (~2.11 Ma, 95% HPD = 1.88–2.35 Ma) in our divergence dating analysis, which corresponds with the extensive uplifts of the QTP (1.7–3.6 Ma) [63]. The main distribution region of sect. *Diphyllon* is in the subtropical (central/east/south) China [6], including the areas between 21° and 34° N of south China. The rapid uplift of the QTP has increased the tectonic activities in Southwest China [64], which may have triggered the early diversification and shaped the geographical pattern of *Epimedium*. In addition, the sect. *Diphyllon* clade (~1.14 Ma, 95% HPD: 1.05–1.23 Ma) was inferred to be slightly older than *E. koreanum* (~0.06 Ma, 95% HPD = 0.02–0.13 Ma). Severe climatic oscillations during Pleistocene glacial-interglacial cycles happened in subtropical China and its adjoining areas [65,66]. The divergence time of sect. *Diphyllon* coincided with the largest glaciation on the QTP that started at approximately 1.2 Mya and reached its maximum between 0.8 and 0.6 Mya [67]. The environmental changes may have further affected the genetic structure and geographical distribution of *Epimedium* species. Thereby, it is likely that the geological events and the climate oscillation of the largest glaciation in the Quaternary contributed to the diversification of *Epimedium* in China.

Evolution of petal characters in Epimedium

Flower morphology is a critical feature that can be used to distinguish *Epimedium* species [1]. According to the size relationship between the petals and inner sepals, *Epimedium* were divided into two groups, namely, large-flowered group (petals longer than inner sepals; >1 cm in diameter) and small-flowered group (petals shorter than inner sepals; <1 cm in diameter) [68,69]. The small-flowered group consists of four petal types, i.e., flat, slightly saccate, saccate, and short spur, whereas the large-flowered group is long spur. Ying proposed that China was the distribution and differentiation center of the genus *Epimedium*, with petals evolving in a continuous way [45]. The flat petal was assumed to be the ancestral type that evolved into four other types, namely, slightly saccate, saccate, short spur, and long spur. However, our ancestral character reconstruction analysis suggested a different history, in which the long spur (large-flowered group) constituted the plesiomorphic state (Fig. 4).

The hypothesis that the evolution of spur length is driven by pollinator shifts has been examined by phylogenetic evidence in recent studies [70,71]. In *Epimedium*, an outcrossing breeding system is prevalent with strong self-incompatibility but high cross-compatibility between infraspecific or interspecific individuals [72,73]. Two types of effective *Epimedium* pollinators, namely, pollen-collecting and nectar-foraging bees, were recognized [74–76]. Species with spurless flowers were pollinated by only the pollen-collecting small bees (*Andrena* spp. and *Lasioglossum* spp.),

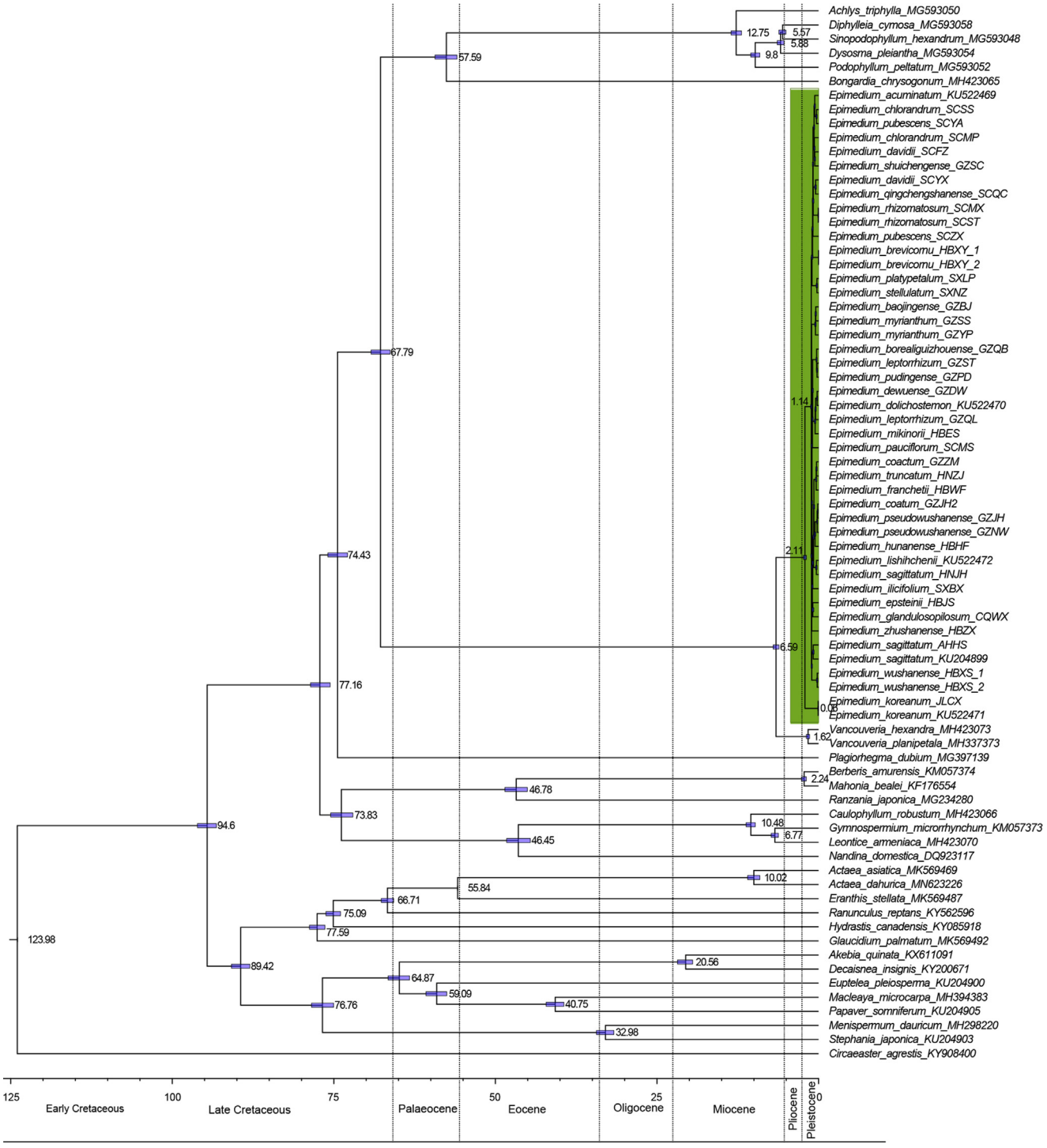


Fig. 5. Beast maximum clade credibility tree inferred from the coding sequence of the 77 protein-coding genes. The divergence times of each clade is displayed near each node. Blue bars represent the 95% highest posterior density for the node ages.

whereas species with short spurs were pollinated by both pollen-collecting visits of *Andrena* spp. and medium sized nectar-foraging visits of *Tetralonia nipponensis* [74,76]. Two types of effective pollinators were also observed in species with long spurs, distinguished by large *Bombus* spp. as nectar-foraging bees [74,75] Suzuki found that pollen-collecting bees foraged successively on the *Epimedium* flowers with different lengths of spurs [74]. The spur length of *Epimedium* had no effect on the foraging behavior of pollen-collecting bees, which might promote gene flow between species with different spur types. As for nectar-foraging bees, *T. nipponensis* that has a short proboscis and *Bombus* spp. that has a long proboscis showed preference for flowers with short and long spurs, respectively, suggesting potential correlations between the spur length of flowers and the proboscis length of bees. The behavior of this type of bees could strengthen the reproductive isolation of *Epimedium* species [74]. Therefore, we suspected that the long spur, which occurs in more than half of the *Epimedium* species, evolved to fit the pollinators shifts. The diversification of petal

shape likely affects the reproductive success and thereby drove the evolution of *Epimedium*. Currently, research about plant-pollinator interactions of *Epimedium* species is still scarce. Apart from adequate investigations about plant–pollinator interaction, other environmental factors should also be considered to better elucidate the evolution of petal characters in *Epimedium* in further studies.

Traditional series-level classification of sect. *Diphyllon* [1] is based on petal traits. However, character optimization indicated that petal shape (scheme 1) is not a suitable and efficient character for series-level classification, because states of this character have evolved multiple times within sect. *Diphyllon* (Fig. S7). Considering the incongruence between molecular phylogeny (Fig. 2) and traditional classification of sect. *Diphyllon* [1], a single morphological character is insufficient to unequivocally circumscribe the main clades in the phylogenetic trees. Thus, combining diagnostic morphological characters for the subdivision of the sect. *Diphyllon* might be a good strategy.

Conclusions

We presented a comparative analysis of 45 plastomes from 32 *Epimedium* species and reported a comprehensive study of their phylogenetic relationships, divergence time estimation, and ancestral character state reconstruction for the first time. Four distinct plastome organizations were found in *Epimedium* according to the variation of IR boundary. The phylogenetic analysis supported the sister relationship of sect. *Macroceras* and sect. *Diphyllon* but could not resolve the further division of sect. *Diphyllon* into series. Considerable inconsistency was observed between the molecular phylogeny and traditional classification of sect. *Diphyllon*. The *Epimedium* ancestors probably diversified in the early Pleistocene (~2.11 Ma). Ancestral state reconstruction showed that *Epimedium* originated from long spur (large-flowered group). These results provide valuable information to elucidate the intricate taxonomy, phylogeny, and evolution process of *Epimedium*.

Compliance with ethics requirements

This article does not contain any studies with human or animal subjects.

CRediT authorship contribution statement

Mengyue Guo: Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Xiaohui Pang:** Conceptualization, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing, Funding acquisition. **Yanqin Xu:** Investigation, Resources. **Wenjun Jiang:** Formal analysis, Data curation. **Baosheng Liao:** Software, Data curation. **Jingsheng Yu:** Formal analysis. **Jiang Xu:** Software, Writing - review & editing. **Jingyuan Song:** Writing - review & editing. **Shilin Chen:** Conceptualization, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge Professor Zhiduan Chen (Institute of Botany, Chinese Academy of Sciences) for valuable comments on an early draft of the manuscript. This work was supported by grants from the National Key Research and Development Program of China (No. 2019YFC1604701) and the National Natural Science Foundation of China (No. 81573541).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2021.06.020.

References

[1] Stearn WT. The Genus *Epimedium* and other Herbaceous Berberidaceae. Portland: Timber Press; 2002.
[2] Ying TS, Boufford DE, Brach AR. Epimedium L. In: Flora of China. Beijing: Science Press & St. Louis: Missouri Botanical Garden Press; 2011. p. 787–99.
[3] Zhang Y, Dang H, Li S, Li J, Wang Y. Five new synonyms in *Epimedium* (Berberidaceae) from China. PhytoKeys 2015;49:1–12.
[4] Zhang Y, Dang H, Li J, Wang Y. The *Epimedium wushanense* (Berberidaceae) species complex, with one new species from Sichuan, China. Phytotaxa 2014;172:39–45.
[5] Zhang Y, Du L, Liu A, Chen J, Wu L, Hu W, et al. The complete chloroplast genome sequences of five *Epimedium* species: lights into phylogenetic and taxonomic analyses. Front Plant Sci 2016;7:306.
[6] Zhang ML, Uhink CH, Kadereit JW. Phylogeny and biogeography of *Epimedium*/*Vancouveria* (Berberidaceae): Western North American-East Asian disjunctions, the origin of European mountain plant taxa, and East Asian species diversity. Syst Bot 2007;32:81–92.
[7] Ma H, He X, Yang Y, Li M, Hao D, Jia Z. The genus *Epimedium*: an ethnopharmacological and phytochemical review. J Ethnopharmacol 2011;134:519–41.
[8] Linnaeus C. Species Plantarum (1957 reprint). London: Ray Society; 1753.
[9] Sun Y, Fung KP, Leung PC, Shaw PC. A phylogenetic analysis of *Epimedium* (Berberidaceae) based on nuclear ribosomal DNA sequences. Mol Phylogenet Evol 2005;35:287–91.
[10] Zhang YJ, Dang HS, Meng AP, Li JQ, Li XD. Karyomorphology of *Epimedium* (Berberidaceae) and its phylogenetic implications. Caryologia 2008;61:223–32.
[11] Sheng MY, Wang LJ, Tian XJ. Karyomorphology of eighteen species of genus *Epimedium* (Berberidaceae) and its phylogenetic implications. Genet Resour Crop Ev 2010;57:1165–76.
[12] Zhang Y, Yang L, Chen J, Sun W, Wang Y. Taxonomic and phylogenetic analysis of *Epimedium* L. based on amplified fragment length polymorphisms. Sci Hortic 2014;170:284–92.
[13] De Smet Y, Goetghebeur P, Wanke S, Asselman P, Samain MS. Additional evidence for recent divergence of Chinese *Epimedium* (Berberidaceae) derived from AFLP, chloroplast and nuclear data supplemented with characterisation of leaflet pubescence. Plant Ecol Evol 2012;145:73–87.
[14] Xu Y, Li Z, Wang Y. Fourteen microsatellite loci for the Chinese medicinal plant *Epimedium sagittatum* and cross-species application in other medicinal species. Mol Ecol Resour 2008;8:640–2.
[15] Guo M, Xu Y, Ren L, He S, Pang X. A systematic study on DNA barcoding of medicinally important genus *Epimedium* L. (Berberidaceae). Genes-Basel 2018;9:637.
[16] Xu Y, Liu L, Liu S, He Y, Li R, Ge F. The taxonomic relevance of flower colour for *Epimedium* (Berberidaceae), with morphological and nomenclatural notes for five species from China. PhytoKeys 2019;118:33–64.
[17] Raubeson LA, Jansen RK. Chloroplast genomes of plants. In: Henry RJ, editor. Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants. Cambridge (MA): CABI Press; 2005. p. 45–68.
[18] Wicke S, Schneeweiss GM, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol 2011;76:273–97.
[19] Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol 2009;7:84.
[20] Niu YT, Jabbour F, Barrett RL, Ye JF, Zhang ZZ, Lu KQ, et al. Combining complete chloroplast genome sequences with target loci data and morphology to resolve species limits in *Triplostegia* (Caprifoliaceae). Mol Phylogenet Evol 2018;129:15–26.
[21] Morales-Briones DF, Liston A, Tank DC. Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus *Lachemilla* (Rosaceae). New Phytol 2018;218:1668–84.
[22] Zhong L, Barrett SCH, Wang XJ, Wu ZK, Sun HY, Li DZ. Phylogenomic analysis reveals multiple evolutionary origins of selfing from outcrossing in a lineage of heterostylous plants. New Phytol 2019;224:1290–303.
[23] Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, et al. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res 2019;47:W65–73.
[24] Lee JH, Kim K, Kim NR, Lee SC, Yang TJ, Kim YD. The complete chloroplast genome of a medicinal plant *Epimedium koreanum* Nakai (Berberidaceae). Mitochondrial DNA A 2016;27:4342–3.
[25] Sun Y, Moore MJ, Zhang S, Soltis PS, Soltis DE, Zhao T, et al. Phylogenomic and structural analyses of 18 complete plastomes across nearly all families of

early-diverging eudicots, including an angiosperm-wide analysis of IR gene content evolution. Mol Phylogenet Evol 2016;96:93–101.
[26] Guo M, Ren L, Xu Y, Liao B, Song J, Li Y, et al. Development of plastid genomic resources for discrimination and classification of *Epimedium wushanense* (Berberidaceae). Int J Mol Sci 2019;20:4003.
[27] Jiang H, Lei R, Ding SW, Zhu S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinf 2014;15:182.
[28] Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. ABySS: a parallel assembler for short read sequence data. Genome Res 2009;19:1117–23.
[29] Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biol 2004;5:R12.
[30] 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:W59–64.
[31] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biol Evol 2013;30:2725–9.
[32] Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res 2004;32:W273–9.
[33] 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.
[34] Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE 2010;5:e11147.
[35] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013;30:772–80.
[36] Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 2009;25:1972–3.
[37] Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015;32:268–74.
[38] 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.
[39] Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 2012;61:539–42.
[40] Nylander JAA. Version 2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University; 2004.
[41] Rambaut A. FigTree version 1.4.2: tree figure drawing tool. 2014. Available from: http://tree.bio.ed.ac.uk/software/figtree.
[42] Zhang YJ, Dang HS, Wang Y, Li XW, Li JQ. A taxonomic revision of unifoliolate Chinese *Epimedium* L. (Berberidaceae). Kew Bull 2011;66:253.
[43] Zhang YJ, Dang HS, Li JQ, Wang Y. Taxonomic notes on three species of *Epimedium* (Berberidaceae) endemic to China. Phytotaxa 2015;204:147–52.
[44] Liu S, Liu L, Huang X, Zhu Y, Xu Y. A taxonomic revision of three Chinese spurless species of genus *Epimedium* L. (Berberidaceae). PhytoKeys 2017;78:23–36.
[45] Ying TS. Petal evolution and distribution patterns of *Epimedium* L. (Berberidaceae). Acta Phytotaxon Sin 2002;40:481–9.
[46] Maddison WP, Maddison DR. Mesquite: A modular system for evolutionary analysis, version 3.6. 2018. Available from: http://mesquiteproject.org.
[47] Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, et al. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput Biol 2014;10:e1003537.
[48] Schorn H. Revision of the Fossil Species of Mahonia from North America. University of California; 1966.
[49] Manchester S. Biogeographical relationships of North American Tertiary floras. Ann Mo Bot Gard 1999;86:472–522.
[50] Magallón S, Gómez-Acevedo S, Sánchez-Reyes LL, Hernóndez-Hernóndez T. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol 2015;207:437–53.
[51] Pigg KB, Devore ML. Paleoactaea gen. nov. (Ranunculaceae) fruits from the Paleogene of North Dakota and the London Clay. Am J Bot 2005;92:1650–9.
[52] von Balthazar M, Pedersen K, Friis E. *Teixeiraea lusitanica*, a new fossil flower from the Early Cretaceous of Portugal with affinities to the Ranunculales. Plant Syst Evol 2005;255:55–75.
[53] Rambaut A, Suchard MA, Xie D, Drummond AJ. Tracer v1.6; 2014. Available from: http://beast.bio.ed.ac.uk/Tracer.
[54] Soltis PS, Folk RA, Soltis DE. Darwin review: angiosperm phylogeny and evolutionary radiations. P Roy Soc B-Biol Sci 2019;286:20190099.
[55] Kane NC, King MG, Barker MS, Raduski A, Karrenberg S, Yatabe Y, et al. Comparative genomic and population genetic analyses indicate highly porous genomes and high levels of gene flow between divergent helianthus species. Evolution 2009;63(8):2061–75.
[56] Alwadani KG, Janes JK, Andrew RL. Chloroplast genome analysis of box-ironbark *Eucalyptus*. Mol Phylogenet Evol 2019;136:76–86.
[57] Del Valle JC, Casimiro-Soriguer I, Buide M, Narbona E, Whittall JB. Whole plastome sequencing within *Silene* section *Psammophilae* reveals mainland hybridization and divergence with the Balearic Island populations. Front Plant Sci 2019;10:1466.
[58] Xu Y, Li Z, Wang Y, Huang H. Allozyme diversity and population genetic structure of three medicinal *Epimedium* species from Hubei. J Genet Genomics 2007;34:56–71.
[59] Dang M, Yue M, Zhang M, Zhao G, Zhao P. Gene introgression among closely related species in sympatric populations: a case study of three walnut (*Juglans*) species. Forests 2019;10:965.
[60] Weng M, Ruhlman TA, Jansen RK. Expansion of inverted repeat does not decrease substitution rates in *Pelargonium* plastid genomes. New Phytol 2017;214:842–51.
[61] Wang W, Chen ZD, Liu Y, Li RQ, Li JH. Phylogenetic and biogeographic diversification of Berberidaceae in the Northern Hemisphere. Syst Bot 2007;32:731–42.
[62] Sun Y, Moore MJ, Landis JB, Lin N, Chen L, Deng T, et al. Plastome phylogenomics of the early-diverging eudicot family Berberidaceae. Mol Phylogenet Evol 2018;128:203–11.
[63] Zhang D, Liu F, Bing J. Eco-environmental effects of the Qinghai-Tibet Plateau uplift during the Quaternary in China. Environ Geol 2000;39:1352–8.
[64] Sun BN, Wu JY, Liu YS, Ding ST, Li XC, Xie SP, et al. Reconstructing Neogene vegetation and climates to infer tectonic uplift in western Yunnan. China. Palaeogeogr Palaeocl 2011;304:328–36.
[65] Axelrod DI, Al-Shehbaz I, Raven PH. History of the modern flora of China. In: Zhang A, Wu S, editors. Floristic Characteristics and Diversity of East Asian Plants. New York: Springer; 1996. p. 43–55.
[66] Qiu YX, Fu CX, Comes HP. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol Phylogenet Evol 2011;59:225–44.
[67] Cui Z, Chen Y, Zhang W, Zhou S, Zhou L, Zhang M, et al. Research history, glacial chronology and origins of Quaternary glaciations in China. Quat Sci 2011;31:749–64.
[68] Stearn WT. New large-flowered Chinese Species of *Epimedium* (Berberidaceae). Curtis's Botan Mag 1993;10:178–81.
[69] Stearn WT. The small-flowered Chinese species of *Epimedium* (Berberidaceae). Kew Bull 1993;48:807–13.
[70] Whittall JB, Hodges SA. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. Nature 2007;447:706–9.
[71] Boberg E, Alexandersson R, Jonsson M, Maad J, Ågren J, Nilsson LA. Pollinator shifts and the evolution of spur length in the moth-pollinated orchid Platanthera bifolia. Ann Bot 2014;113:267–75.
[72] Suzuki K. Breeding system and crossability in Japanese *Epimedium* (Berberidaceae). J Plant Res 1983;96:343–50.
[73] Sheng M, Chen Q, Wang L, Tian X. Hybridization among *Epimedium* (Berberidaceae) species native to China. Sci Hortic 2011;128:342–51.
[74] Suzuki K. Pollination system and its significance on isolation and hybridization in Japanese *Epimedium* (Berberidaceae). J Plant Res 1984;97:381–96.
[75] Li YX, Quan QM, Sun GL. Effect of floral morphology on fruit set in *Epimedium sagittatum* (Berberidaceae). Plant Syst Evol 2009;279:51–8.
[76] Quan Q, Li Y. Effect of morphs on reproductive biology of *Epimedium pubescens* (Berberidaceae): a species endemic to China. Int J Agric Biol 2018;20:945–50.