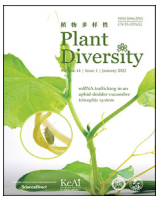
Plant Diversity 44 (2022) 70–82

Contents lists available at ScienceDirect

Plant Diversity

journal homepage: http://www.keaipublishing.com/en/journals/plant-diversity/
http://journal.kib.ac.cn



Research paper

Comparative chloroplast genome analysis of medicinally important *Veratrum* (Melanthiaceae) in China: Insights into genomic characterization and phylogenetic relationships

Ying-Min Zhang [a, 1], Li-Jun Han [a, 1], Cong-Wei Yang [a], Zi-Li Yin [b], Xing Tian [a], Zi-Gang Qian [a, b, **], Guo-Dong Li [a, b, *]

[a] Faculty of Traditional Chinese Pharmacy, Yunnan University of Chinese Medicine, Kunming 650500, Yunnan, China
[b] Yunnan Key Laboratory for Dai and Yi Medicines, Yunnan University of Chinese Medicine, Kunming 650500, Yunnan, China

ARTICLE INFO

Article history:
Received 7 January 2021
Received in revised form
19 May 2021
Accepted 21 May 2021
Available online 1 June 2021

Keywords:
Veratrum
Melanthiaceae
Chloroplast genome
Comparative analysis
Phylogeny

ABSTRACT

Members of *Veratrum* are perennial herbs widely used in traditional Chinese medicine to induce vomiting, resolve blood stasis and relieve pain. However, the intrageneric classification and phylogenetic relationships within *Veratrum* have long been controversial due to the complexity of morphological variations and lack of high-resolution molecular markers. In this study, we reevaluated the infrageneric relationships with the genus *Veratrum* using complete chloroplast genome sequence data. Herein, the complete cp genomes of ten species of *Veratrum* were newly sequenced and characterized. The complete cp genomes of ten species of *Veratrum* had the typical quadripartite structure, ranging from 151,597 bp to 153,711 bp in size and comprising a total of 135 genes. The structure of *Veratrum* cp genomes (i.e., gene order, content, and genome components) was highly similar across species. The number of simple sequence repeats (SSRs) ranged from 63 to 78, and of long repeats ranged from 31 to 35. Eight highly divergent regions (*ndhF*, *psbC-psbZ*, *psbK-psbI*, *rpoB-trnC*_GCA, *trnK*_UUU-*trnQ*_UUG, *trnS*_GCU-*trnG*_UCC, *trnT*_UGU-*trnL*_UAA and *ycf1*) were identified and are potentially useful for the DNA barcoding of *Veratrum*. Phylogenetic analysis among 29 taxa based on cp genomes, total genes, protein-coding genes and intergenic regions strongly supported the monophyly of *Veratrum*. The circumscription and relationships of the infrageneric taxa of *Veratrum* were well-presented with great resolution. These results will facilitate the identification, taxonomy, and utilization of *Veratrum* plants as well as the evolutionary studies of Melanthiaceae.

1. Introduction

The genus *Veratrum* L. is considered one of the most important groups in the family Melanthiaceae (Liliflorae), with approximately 17–45 species of perennial herbaceous plants (Tsi, 1980; Chen and Takahashi, 2000; Zomlefer et al., 2001, 2003; Chase et al., 2016). Although members of *Veratrum* are widely distributed from the temperate to arctic zones of the Northern Hemisphere, the majority are native to the Eastern Asia region (Zomlefer et al., 2001; Liao et al., 2007). In China, there are roughly 13 species and one variety of *Veratrum*, several of which have been used medically for over 1700 years, including *V. nigrum* L., *V. schindleri* O. Loes, *V. maackii* Regel and others (Tsi, 1980; Wu et al., 1999; Chen and Takahashi, 2000). For instance, the dried roots and rhizomes of *V. nigrum* (*lilu* in Chinese) have been used to treat aphasia arising from apoplexy, “wind-type dysentery”, scabies, jaundice, and chronic malaria (Wu et al., 1999). Other species, such as *V. grandiflorum* (Maxim. ex Miq.) O. Loes., *V. mengtzeanum* O. Loes., *V. stenophyllum* Diels, and *V. taliense* O. Loes., are sourced as the folk medicine “*pimacao*” in China to treat bruises, rheumatic pain, and

* Corresponding author. Faculty of Traditional Chinese Pharmacy, Yunnan University of Chinese Medicine, Kunming 650500, Yunnan. China.
** Corresponding author. Faculty of Traditional Chinese Pharmacy, Yunnan University of Chinese Medicine, Kunming 650500, Yunnan, China.
E-mail addresses: qianzig@aliyun.com (Z.-G. Qian), gammar116@163.com (G.-D. Li).
Peer review under responsibility of Editorial Office of Plant Diversity.
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.pld.2021.05.004

wound hemostasis (Wang et al., 2005). These plants are also currently important ingredients in commercialized modern drugs (Yin et al., 2014; Li et al., 2016). The major active ingredient isolated from the roots of each of these *Veratrum* species is several steroidal alkaloid, which is known to decrease blood pressure, prevent platelet aggregation and thrombosis, as well as act as an anti-inflammatory, analgesic and antitumor agent (Turner et al., 2019; Cong et al., 2020). However, *Veratrum* plants also contain toxic compounds such as cevanine alkaloids (Kikkawa et al., 2017; Turner et al., 2019). These compounds cause nausea, bradycardia, hypotension, and apnea shortly after ingestion, and in some cases result in death (Chandler and McDougal, 2014). It has been noted that *Veratrum*-related medicinal products are often adulterated with several closely related species, thus undermining food security and pharmacological efficacy (Kikkawa et al., 2017). Hence, utilization of *Veratrum* in medicinal materials requires accurate identification techniques.

Taxonomic classification and the phylogenetic relationships within *Veratrum* have been controversial (Zimmerman, 1958; Kupchan et al., 1961; Bodkin, 1978; Zomlefer, 1997; Zomlefer et al., 2003). Over the years, different subgenera, sections, and subsections of *Veratrum* have been proposed (Linnaeus, 1753; Loesener, 1926, 1927, 1928; Nakai, 1937a, b; Kupchan et al., 1961; Bodkin, 1978; Tamura, 1998). Currently, researchers accept two sections (section *Veratrum* and section *Fuscoveratrum*) and section *Fuscoveratrum* were divided into two subsections (subsection *Pseudoanti* and subsection *Asiaveratrum*) (Zomlefer et al., 2003). Although a number of phylogenetic studies based on single and multi-loci DNA sequences have been conducted (Liao et al., 2007; Kim et al., 2014, 2016), the interspecific relationships in *Veratrum* still remain ambiguous for some taxa. These include the identity of *V. japonicum* (Baker) Loes. f. plants from China, which have been re-assessed in *Flora of China* and identified as *V. schindleri* (Chen and Takahashi, 2000), and *V. japonicum* outside of China, which The Plant List reduced to a variety of *V. maackii* (http://www.theplantlist.org/). Clarifying these uncertainties in *Veratrum* phylogeny will likely require high resolution molecular evidence.

The chloroplast (cp) genome has a relatively conserved structure in higher plants, consisting of a double-stranded circular molecule that is 120–160 kb in length. The cp genome comprises a large single-copy (LSC) region and a small single-copy (SSC) region, separated by two identical copies of inverted repeat (IR) regions (Moore et al., 2010). The cp genome has been widely used for evolutionary, taxonomic and species diversity-related studies due to its highly conserved genome structure, maternal inheritance, low to moderate evolutionary rate, and low effective population sizes features (Sugiura, 1992; Moore et al., 2010). Recent studies have demonstrated that cp genome sequences are remarkable at resolving phylogenetic relationships at different taxonomic levels (Yang et al., 2018, 2019; Gu et al., 2019; Xie et al., 2019). However, despite this acknowledged utility, little information on *Veratrum* cp genome sequences has been published.

In this study, we aimed to (1) identify highly variable regions within *Veratrum* chloroplast genomes that can be used as DNA barcodes and (2) reconstruct the phylogenetic relationships of *Veratrum*. For these purposes, we sequenced, assembled, and annotated the complete chloroplast genomes of ten species of *Veratrum* from China, and then characterized the cp genomes of these species, examining variations in simple sequence repeats (SSRs), long repeats and other highly variable regions. For phylogenetic analysis of *Veratrum*, we used four sets of sequence data: complete cp genomes, total genes, protein-coding genes, and intergenic regions.

2. Materials and methods

2.1. Plant materials and DNA extraction

Fresh leaves were collected in the field (Fujian, Liaoning, Sichuan, and Yunnan provinces, China; see Table S1) from ten *Veratrum* species: *V. dahuricum* (Turcz.) O. Loes., *V. grandiflorum*, *V. japonicum*, *V. maackii*, *V. nanchuanense* S. Z. Chen et G. J. Xu, *V. nigrum*, *V. oblongum* O. Loes., *V. schindleri*, *V. stenophyllum*, and *V. taliense*. Leaves were stored in allochroic silica gels prior to be transported back to the laboratory. Voucher specimens were prepared and deposited in the Herbarium of Yunnan University of Chinese Medicine, China. Note that *V. japonicum* of Chinese origin has been previously treated as *V. schindleri* (Chen and Takahashi, 2000); to aid our discussion, we retained the taxa as the former.

Total genomic DNA extraction was carried out using a modified Cetyltrimethylammonium Bromide (CTAB) protocol (Doyle et al., 1992). The quality and quantity of the DNA extracts were determined using both the 1.0% agarose gel electrophoresis and the Nanodrop™ 2000 spectrophotometer (Thermo Fisher Scientific, USA).

2.2. Genome sequencing, assembly and annotation

Paired-end libraries (insert size 350 bp) were constructed prior to next-generation sequencing on an Illumina HiSeq 2500-PE platform (Illumina, United States). High quality reads were obtained by filtering the low-quality reads and connector sequences with the NGS QC Toolkit set at default parameters (Patel and Jain, 2012). Filtered reads were assembled *de novo* using NOVOPlasty (Dierckxsens et al., 2017) by mapping against the reference cp genome of *Veratrum oxysepalum* Turcz. (NC_022715; Do et al., 2013). Gene annotation was performed with the online annotation tool DOGMA (Wyman et al., 2004) and subsequently corrected using Geneious Prime® 2020.1.1 (Kearse et al., 2012). The cp genome map was illustrated with OGDRAW v.1.3.1 (Greiner et al., 2019). All complete cp genomes were deposited into GenBank under the accession numbers MN604405, MN613588–MN613595, MN699635 (Table 1).

2.3. Codon usage, SSRs and long repeat analysis

Relative synonymous codon usage (RSCU) of 84 protein-coding genes was assessed using MEGA X with default parameters (Kumar et al., 2018). SSRs were detected using IMEx (Mudunuri and Nagarajaram, 2007) with the search parameters set as 10, 5, 4, 3, 3, and 3 for mono-, di, tri-, tetra-, penta- and hexa-nucleotides, respectively. The REPuter software (Kurtz et al., 2001) was used for the identification of palindromic, forward, reverse and complement repeats present in the cp genome, in which the Hamming distance was set as 3 and the minimum repeat size was 30 bp.

2.4. Genome comparison and structural analysis

The cp genomes were aligned and visualized with mVISTA under the Shuffle-LAGAN mode (Frazer et al., 2004). Prior to checking the possible rearrangements in the cp genomes using MAUVE, the cp genome sequences were aligned using progresssiveMauve (Kurtz et al., 2004). The gene content adjacent to the borders of the two single copies and the IR regions in the cp genomes were visualized and compared using IRscope (Amiryousefi et al., 2018). The nucleotide variability across the cp genome sequences were

Table 1
Summary of complete chloroplast genomes of twelve *Veratrum* species.

Genome Characteristic	Accession number	Length (bp)				Gene number				GC cotent (%)			
		Full	LSC	SSC	IR	Full	PCG	tRNA	rRNA	Full	LSC	SSC	IR
V. dahuricum	MN699635	153,688	83,363	17,607	26,359	135	83	38	8	37.7	35.7	31.4	42.9
V. grandiflorum	MN613592	153,711	83,367	17,628	26,358	135	83	38	8	37.7	35.7	31.4	42.9
V. japonicum	MN613594	151,842	81,885	17,533	26,212	135	84	38	8	37.7	35.8	31.3	42.9
V. maackii	MN613590	151,597	81,822	17,473	26,151	135	84	38	8	37.7	35.7	31.4	42.9
V. nanchuanense	MN613591	153,692	83,368	17,608	26,358	135	83	38	8	37.7	35.7	31.4	42.9
V. nigrum	MN613595	151,599	81,823	17,474	26,151	135	84	38	8	37.7	35.7	31.4	42.9
V. oblongum	MN613593	151,767	81,997	17,480	26,145	135	84	38	8	37.7	35.7	31.4	42.9
V. schindleri	MN613588	151,768	81,862	17,530	26,188	135	84	38	8	37.7	35.8	31.3	42.9
V. stenophyllum	MN613589	152,028	82,120	17,558	26,175	135	84	38	8	37.8	35.8	31.4	43.0
V. taliense	MN604405	152,040	82,122	17,558	26,180	135	84	38	8	37.8	35.8	31.5	43.0
V. mengtzeanum	MN589932	152,051	82,111	17,544	26,198	135	84	38	8	37.8	35.8	31.5	42.9
V. oxysepalum	NC_022715	153,699	83,372	17,607	26,360	135	83	38	8	37.7	35.7	31.4	42.9

New sequences from this study are in boldface.

analyzed using DnaSP v.6.0 (Rozas et al., 2017), with a window length of 600 bp and a step size of 200 bp.

2.5. Phylogenetic analysis

Phylogenetic reconstruction was carried out by including 27 taxa from the family Melanthiaceae, while two species, *Lilium henryi* Baker (NC_035570) and *Smilax china* L. (HM536959) were selected as outgroup. Phylogenetic analyses were conducted on four different data sets, each analyzed separately: (1) complete cp genome sequences (data set-1); (2) a combination of total gene sequences, which included the protein-coding genes, transfer RNAs (tRNAs), and ribosomal RNAs (rRNAs) (data set-2); (3) a combination of protein-coding gene sequences (data set-3); and (4) a combination of intergenic region sequences (data set-4). For the data sets other than the complete cp genome sequence, targeted sequences were extracted and concatenated using Geneious Prime® 2020.1.1 (Kearse et al., 2012). All data sets were aligned using MAFFT v.7.388 based on default parameters (Katoh and Standley, 2013) and phylogenetic analysis was conducted using both the maximum likelihood (ML) and Bayesian inference (BI) methods. The best-fit evolutionary model for data sets 1, 2, and 3 was the generalized-time reversible (GTR) with invariants (+I) and discrete Gamma (+G) (GTR + I + G) model, as determined by the Akaike information criterion (AIC) using ModelFinder (Kalyaanamoorthy et al., 2017); for the concatenated data set-4, the optimal model was GTR + G. The ML tree was constructed using RAxML v.8.2.4 (Stamatakis, 2014), applied with 1000 bootstrap replicates at each branch node. BI analysis was conducted using MrBayes v.3.2.6 (Huelsenbeck and Ronquist, 2001). The Markov chain Monte Carlo (MCMC) algorithm was set for 2,000,000 generations, and trees were sampled every 1000 generations. The first 25% of the trees were discarded as burn-in, while the remaining trees were used to generate the consensus tree.

3. Results

3.1. General features of Veratrum chloroplast genomes

The cp genomes exhibited a typical circular quadripartite structure and had highly similar gene order and genomic structure. The sizes of cp genomes ranged from 151,597 bp (*Veratrum maackii*) to 153,711 bp (*V. grandiflorum*) in length (Table 1, Fig. 1). The LSC region ranged from 81,822–83,372 bp and the SSC region ranged between 17,473–17,628 bp; in all cp genomes examined, these two regions are separated by two IR regions (26,145–26,360 bp). In all ten *Veratrum* species, the overall GC content of the cp genomes was between 37.7–37.8%. Moreover, the GC content of the IRs (42.9–43.0%) was higher than that of the LSC (35.7–35.8%) and SSC (31.3–31.5%) regions.

The majority of these cp genomes contained 135 genes, including 84 protein-coding, 38 tRNA, and eight rRNA genes (Table 1); however, in three species (*Veratrum dahuricum*, *V. grandiflorum*, and *V. nanchuanense*) one protein-coding gene, the *infA* (transition initiation factor 1), was absent. The LSC region of *Veratrum* contains 61–62 protein-coding genes and 21 tRNA genes, whereas the SSC region contains 12 protein-coding genes and one tRNA gene. Five protein-coding genes and eight tRNA genes were duplicated in the IR regions. Of these genes, 17 contain intronic regions, including eight protein-coding genes (*atpA*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, and *rpoC1*), six tRNA genes (*trnA*_UGC, *trnG*_UCC, *trnI*_GAU, *trnK*_UUU, *trnL*_UAA and *trnV*_UAC), which contain a single intron, and three protein-coding genes (*clpP*, *rps12*, and *ycf3*), which contain two introns (Table 2). The *rps12* gene was identified as a unique trans-spliced gene with a pair of duplicated 3′-end exons in the IR regions and one 5′-end exon in the LSC region; both the *ycf15* and the *ycf68* genes contain a number of internal stop codons in the cp genome.

3.2. Codon usage analysis

The protein-coding genes of the cp plastomes of *Veratrum* species encoded between 26,072 (*V. dahuricum*) to 26,154 (*V. oblongum*) codons (Table S2). The most abundant codons encoded leucine (10.3–10.4% in each species); whereas the second least abundant codons encoded cysteine (1.2% in each species); the least abundant codon is the stop codon, which consists of only 83–85 codons (0.3% in each species) (Fig. 2).

Random synonymous codon usage values showed that codons with A/U-ends were most preferred among amino acids (Table S2). The highest values of random synonymous codon usage were recorded for arginine-encoded AGA (1.937–1.982), whereas the lowest values were recorded for serine-encoded AGC (0.325–0.339). Codon bias (RSCU = 1) was not detected in methionine-encoded AUA or tryptophan-encoded UGG (Fig. 3).

3.3. SSRs and long repeats analysis

The number of SSRs in the complete cp genome sequences ranged from 63 (*Veratrum taliense*) to 78 (*V. oblongum*) (Fig. 4a, Table S3). The most common types of SSRs were the mononucleotide repeats (64.5%), followed by the dinucleotide (18.0%), the tetranucleotide repeats (7.9%), and the trinucleotide repeats (5.3%). Hexanucleotide repeats were only detected in the cp

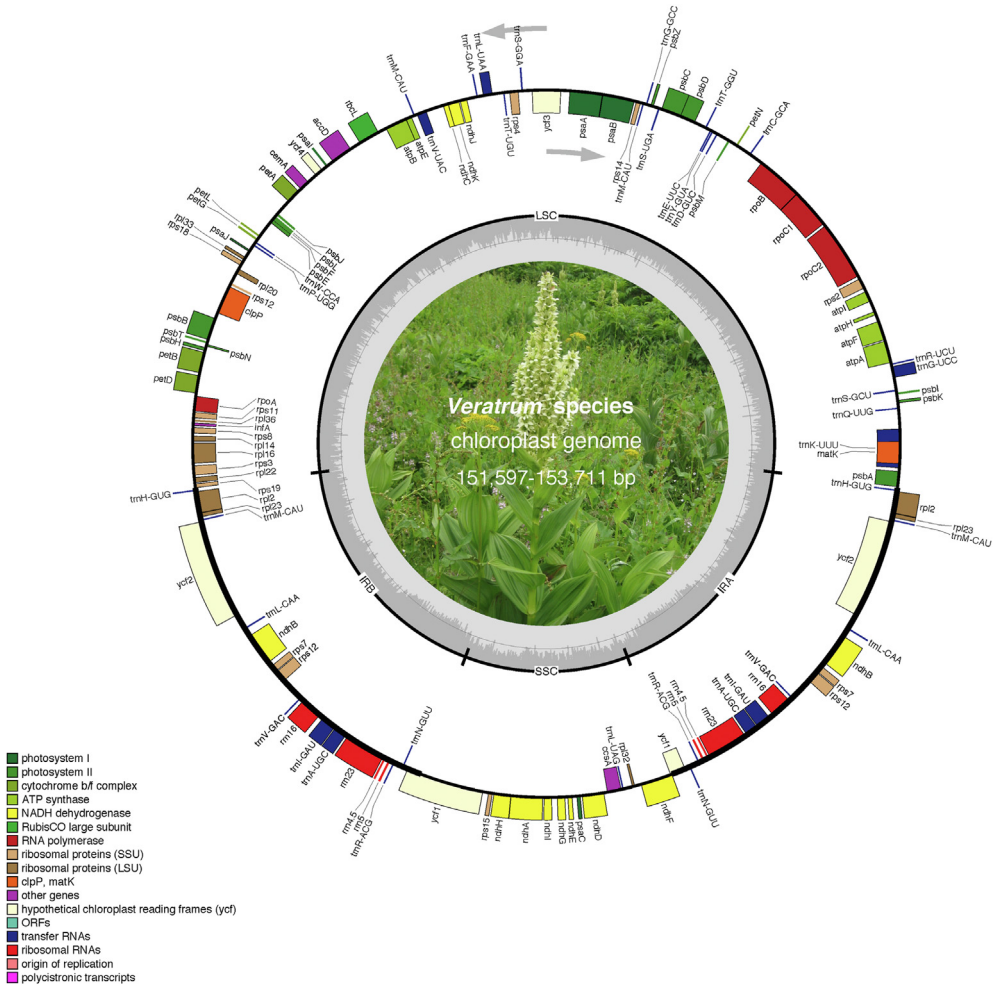


Fig. 1. Gene maps of twelve *Veratrum* species. The colored bars indicate different functional groups. Thick lines indicate the extent of the inverted repeat regions (IRa and IRb), which separate the genome into small (SSC) and large (LSC) single copy regions. Genes drawn inside the circle are transcribed clockwise, while those outside of the circle are transcribed counter-clockwise. The dark gray inner circle corresponds to GC content, the light-gray to AT content.

genome sequences of *V. japonicum*, *V. maackii*, *V. nigrum*, and *V. schindleri* (Fig. 4a). A total of three, two, five, eight, and seven types were recorded for the mono-, di-, tri-, tetra-, penta-, and hexa-nucleotides SSRs, respectively (Fig. 4b). The majority (96%) of these SSRs were located in single-copy regions, while few were recorded in the inverted regions (4%) (Fig. 4c). SSRs were more abundant in intergenic region (57.5%) that in coding regions (23.9%) or introns (19.5%) (Fig. 4d).

A total of 403 long repeats were detected in the complete cp genome sequences, including 150 forward, 226 palindromic and 27 reverse repeats (Fig. 5a, Table S4). The repeat numbers varied from 31 (*Veratrum oblongum*) to 35 (*V. dahuricum*, *V. grandiflorum*, *V. nanchuanense* and *V. oxysepalum*) (Fig. 5a). Long repeats range from 30 to 41 bp in length, whereas only one palindromic repeat was recorded with a length more than 54 bp (e.g., *V. dahuricum*, *V. grandiflorum*, *V. nanchuanense*, and *V. oxysepalum*) (Fig. 5b). Most of the long repeats are located in the LSC region, while a small portion of repeats are distributed along the SSC and IR regions (Fig. 5c). In the LSC region, 38.9% of the repeats are in the intergenic region, while others are dispersed in coding regions (34.3%) and introns (26.8%) (Fig. 5d).

3.4. Comparative analysis of cp genome structures

The visualization analysis of the alignment using mVISTA program and MAUVE program showed that the genomic order and orientation were highly conserved without observed rearrangements, except for slight variations in size and gene positioning (Figs. 6 and 7). Notably, IR regions exhibited less divergence than SSC and LSC regions (Figs. 6 and 7). The non-coding region was more variable than the coding region; the most highly divergent regions in the twelve cp genomes appeared in the intergenic regions, such as *ndhF-rpl32*, *psbC-psbZ*, *rps7-trnV*_GAC, *trnK*_UUU-*trnQ*_UUG and *trnS*_GCU-*trnG*_UCC (Fig. 6). Among coding regions, *petD*, *rpl22*, and *ycf1* genes were relatively divergent (Fig. 6).

When we compared the LSC-IRb, IRb-SSC, SSC-IRa, and IRa-LSC junctions of the twelve *Veratrum* cp genomes, we found that the larger cp genomes of four species (*V. dahuricum*, *V. grandiflorum*, *V. nanchuanense* and *V. oxysepalum*) were due to increased IR lengths (Fig. 8). The LSC-IR and SSC-IR borders of the twelve cp genomes contain six genes (*ndhF*, *rpl2*, *rpl22*, *rps19*, *trnH* and *ycf1*). The *rps19* gene is 2 bp away from the IRb-LSC junction (JLB) in *V. dahuricum*, *V.*

Table 2
Genes in the chloroplast genome of *Veratrum* species.

Gene Category	Name of Genes
Large subunit of ribosomal	*rpl2*[a,c]、*14*、*16*[a]、*20*、*22*、*23*[c]、*32*、*33*、*36*
Small subunit of ribosomal	*rps 2*、*3*、*4*、*7*[c]、*8*、*11*、*12*[b,c,d]、*14*、*15*、*18*、*19*
DNA dependent RNA polymerase	*rpoA*、*rpoB*、*rpoC1*[a]、*rpoC2*
rRNA genes	*rrn4.5*[c]、*rrn5*[c]、*rrn16*[c]、*rrn23*[c]
tRNA genes	*trnA*_UGC[a,c]、*trnC*_GCA、*trnD*_GUC、*trnE*_UUC、*trnF*_GAA、*trnfM*_CAU、*trnG*_GCC、*trnG*-UCC[a]、*trnH*_GUG[c]、*trnI*_CAU[c]、*trnI*_GAU[a,c]、*trnK*_UUU[a]、*trnL*_CAA[c]、*trnL*-UAA[a]、*trnL*_UAG、*trnM*-CAU、*trnN*_GUU[c]、*trnP*_UGG、*trnQ*_UUG、*trnR*_ACG[c]、*trnR*_UCU、*trnS*_GCU、*trnS*_GGA、*trnS*_UGA、*trnT*_GGU、*trnT*_UGU、*trnV*_GAC[c]、*trnV*_UAC[a]、*trnW*_CCA、*trnY*_GUA
Photosystem I	*psaA*、*B*、*C*、*I*、*J*
Photosystem II	*psbA*、*B*、*C*、*D*、*E*、*F*、*H*、*I*、*J*、*K*、*L*、*M*、*N*、*T*、*Z*
NADH oxidoreductase	*ndhA*[a]、*B*[a,c]、*C*、*D*、*E*、*F*、*G*、*H*、*I*、*J*、*K*
Cytochrome *b*6/f complex	*petA*、*B*[a]、*D*[a]、*G*、*L*、*N*
ATP synthase	*atpA*[a]、*B*、*E*、*F*、*H*、*I*
Rubisco	*rbcL*
Maturase	*matK*
Translational	*infA*
Protease	*clpP*[b]
Envelop membrane protein	*cemA*
Subunit of acetyl-CoA	*accD*
c-type cytochrome synthesis gene	*ccsA*
Conserved open reading	*ycf1*、*ycf2*[c]、*ycf3*[b]、*ycf4*、*ycf15*[c]、*ycf68*[c]

[a] Gene containing a single intron.
[b] Gene containing two introns.
[c] Gene with two copies.
[d] Trans-splicing gene.

grandiflorum, *V. nanchuanense* and *V. oxysepalum*, whereas it crosses the IRb-SSC junction (JSB) in the other species. The IRb-SSC junction (JSB) is located in the Ψ*ycf1* region in the chloroplast genomes of all *Veratrum* species and extends a different length (1–10 bp) into the SSC region in all genomes; the IRb region includes 960 bp to 992 bp of the Ψ*ycf1* gene. In contrast, *ycf1* is mainly located in the SSC region, ranging from 4394 bp to 4421 bp, and another 960 bp to 992 bp into the IRa region. In addition, the *trnH* gene is located in the IRa region, 140–143 bp away from the IRa-LSC junction (JSA) in the twelve *Veratrum* chloroplast genomes species.

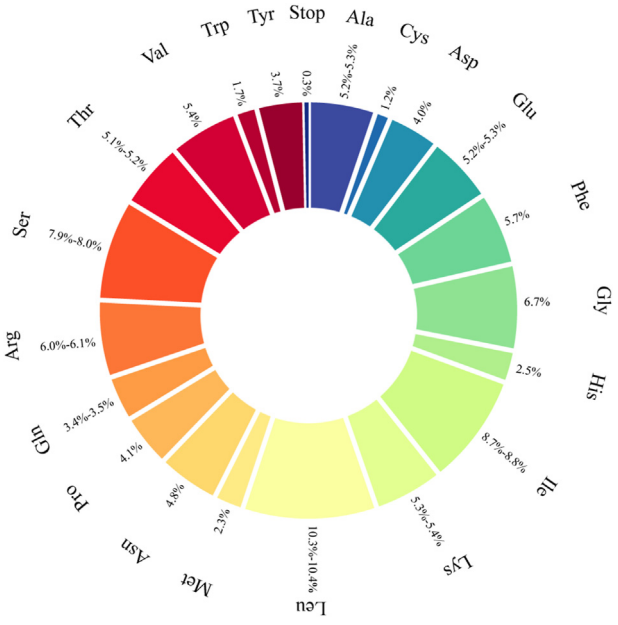


Fig. 2. Codon content for 21 amino acids and stop codons in all protein-coding genes of twelve *Veratrum* CP genome.

3.5. Nucleotide diversity

The cp genomes of *Veratrum* have eight highly variable regions (Fig. 9). All highly divergent fragments were found in the two single-copy regions (Pi = 0.01469), whereas no highly variable loci were detected in the IR regions (Pi = 0.00218). Highly variable regions include the intergenic regions *psbC-psbZ*, *psbK-psbI*, *rpoB-trnC*_GCA, *trnK*_UUU-*trnQ*_UUG, *trnS*_GCU-*trnG*_UCC and *trnT*_UGU-*trnL*_UAA as well as two protein-coding genes, *ndhF* and *ycf1*. The *psbC-psbZ* region had the highest nucleotide diversity (Pi = 0.03576), followed by *rpoB-trnC*_GCA (Pi = 0.03477), *trnT*_UGU-*trnL*_UAA (Pi = 0.03278), *trnK*_UUU-*trnQ*_UUG (Pi = 0.03073), *psbK-psbI* (Pi = 0.02914), *ycf1* (Pi = 0.02896), *trnS*_GCU-*trnG*_UCC (Pi = 0.02843), and *ndhF* (Pi = 0.02694) (Table S5).

3.6. Phylogenetic relationship

The ML and BI trees based on four independent data sets (i.e., complete cp genome, total genes, protein-coding genes, and intergenic region) showed similar topologies (Figs. 10 and 11). However, the branch node support across ML and BI trees (BS ≥ 98% and PP = 1) was higher when based on cp genome sequences than when based on total gene, protein-coding gene, or intergenic region sequence data.

Our trees indicate that Melanthiaceae is monophyletic with five distinct tribes: Chionographideae, Heloniadeae, Melanthieae, Parideae (Trillieae), and Xerophylleae. Melanthieae forms an independent clade that was identified as the basal group of the family tree. For the *Veratrum* clade, the divergence at the first branch node divided the 13 taxa into two groups, consisting of four and nine taxa each, which corresponds to the two sections—*Veratrum* and *Fuscoveratrum* (Fig. 10). Section *Fuscoveratrum* is further divided into two groups with three and six taxa each, which correspond to the formation of two subsections—*Pseudoanti* and *Asiaveratrum*. The *V. japonicum* of Chinese origin clustered with *V. schindleri*, while both taxa form a clade with another *V. japonicum*, which is sister to the *V. oblongum* + *V.*

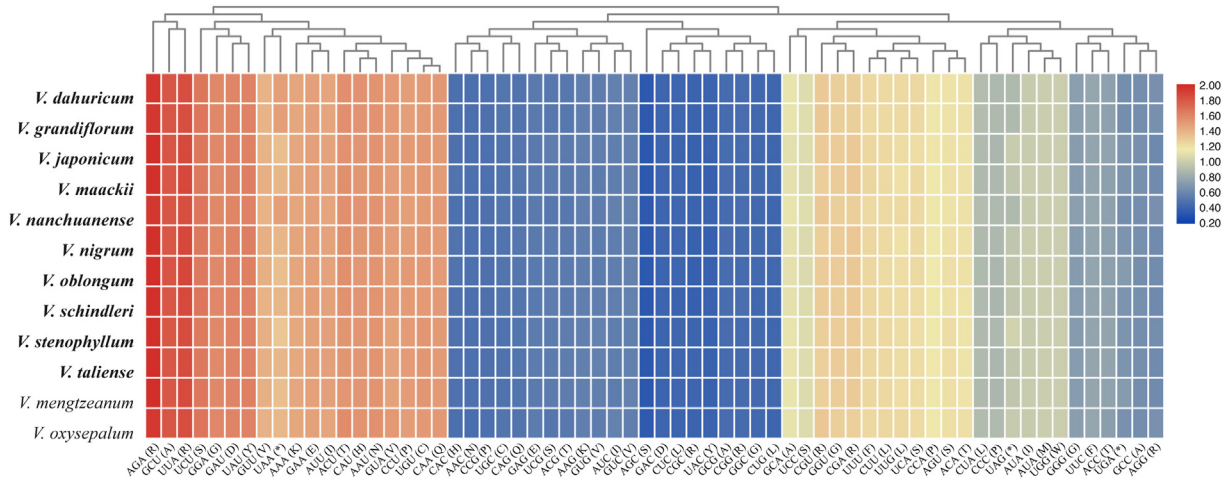


Fig. 3. Heatmap analysis of relative synonymous codon usage (RSCU) values among the twelve *Veratrum* species. Color key: higher red values indicate higher RSCU values, and lower blue values indicate lower RSCU values.

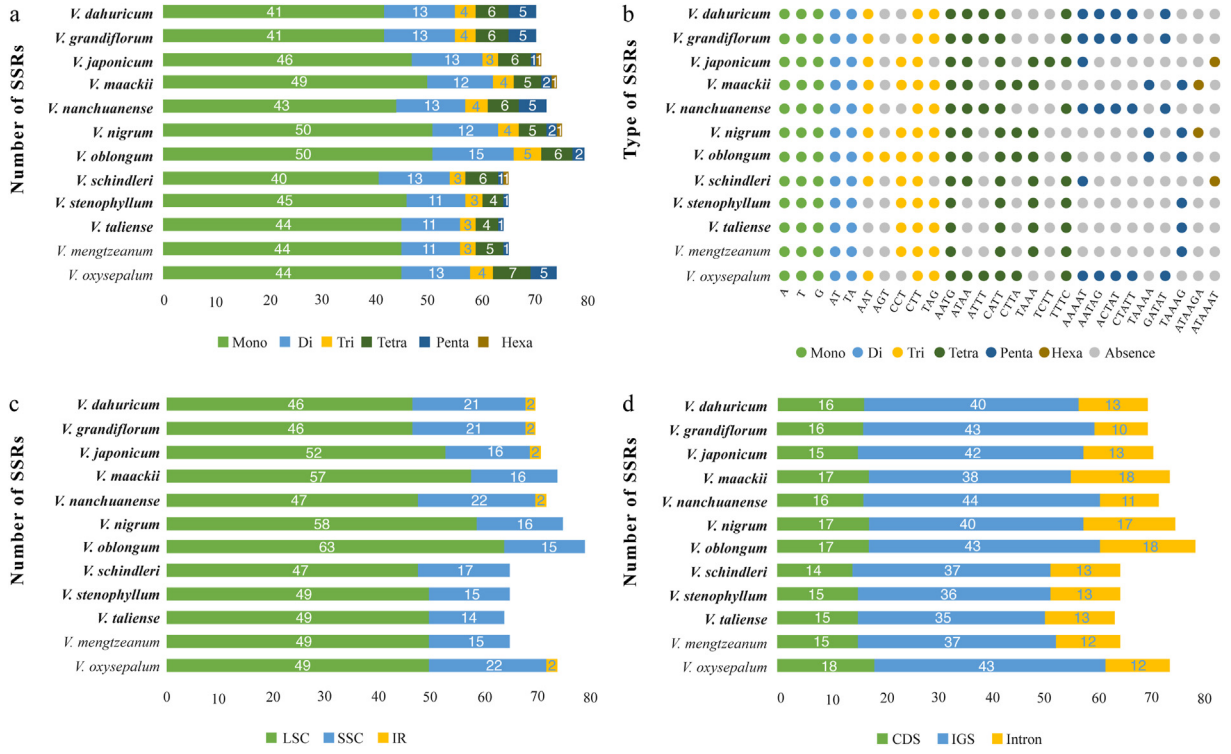


Fig. 4. Analysis of simple repeat sequences (SSR) in twelve *Veratrum* cp genomes. (a) Number of six SSRs repeat. (b) Type of shared SSRs among the twelve cp genomes. (c) Number of SSRs in LSC, SSC, and IR regions. (d) Number of SSRs in the coding regions (CDS), intergenic region (IGS), and introns.

maackii + *V. nigrum* clade. *V. maackii* remains sister to *V. nigrum*, while *V. oblongum* diverges before them; *V. maackii* is not closely-related to *V. japonicum* or *V. schindleri* when compared to *V. oblongum*.

4. Discussion

4.1. Chloroplast genome features

In general, cp genomes include a pair of large inverted repeats that each may reach an average size of 25k bp (Aii et al., 1997). The structural integrity of the complete cp genome is highly affected by the structure of the IR; changes in the structure of the cp genome are often associated with expansions and contractions of the IR regions (Kim and Lee, 2004). The IR regions of cp genomes vary from 10k to 76k bp. Largely due to the evolutionary transfer of complete genes from the SSC regions into the IR or vice versa (Chumley et al., 2006; Wicke et al., 2011). Although we detected small variations around the IR borders, the IR regions of cp genomes used in this study showed only a modest expansion, similar to that found in most land plants (Wolf et al., 2010). IR border variation in

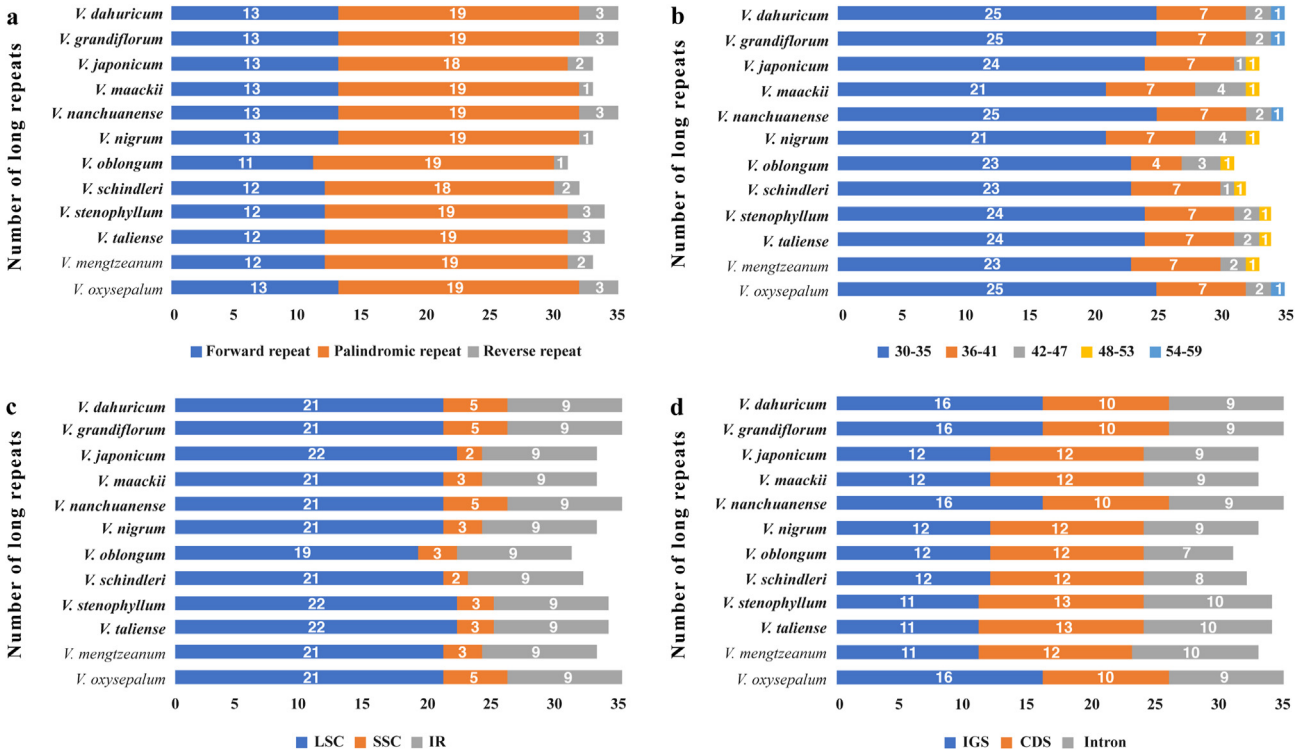


Fig. 5. Analysis of long repeats in twelve *Veratrum* cp genomes. (a) Number of forward, palindromic, and reverse long repeats. (b) Number of long repeats by length. (c) Number of long repeats in LSC, SSC, and IR. (d) Number of long repeats in the coding regions (CDS), intergenic region (IGS), and introns.

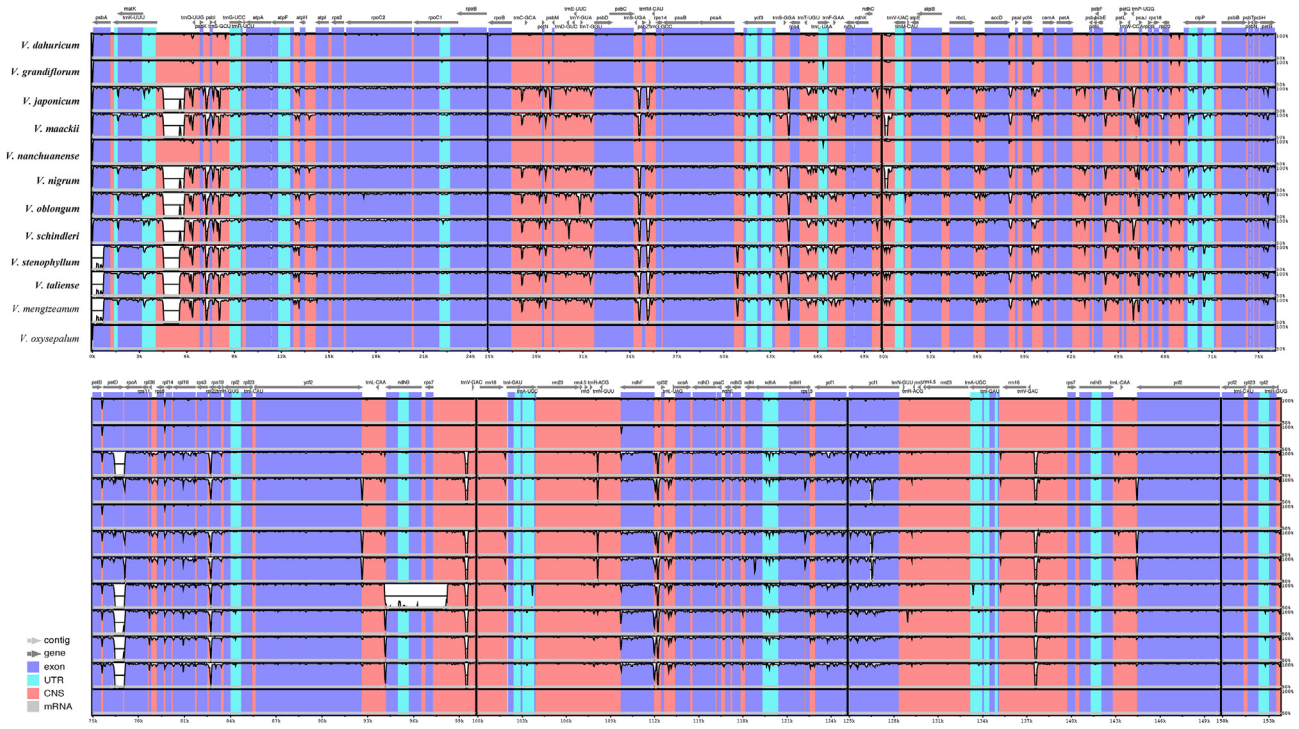


Fig. 6. Comparison of chloroplast genomes of twelve *Veratrum* species. Gray arrows above the alignment indicate the direction of the gene. The dark blue regions represent exons, the light-blue regions represent untranslated regions (UTRs), and the pink regions represent conserved non-coding sequences (CNS). The y-axis represents the percent identity ranging from 50% to 100%.

Fig. 7. Mauve multiple alignment of twelve *Veratrum* cp genomes, with *Veratrum oxysepalum* as the reference.

Veratrum cp genomes is likely consistent with that of other species within the family.

The chloroplast genomes of most *Veratrum* species examined contained the same numbers of protein-coding, tRNA, and rRNA genes. However, the chloroplast genomes of three species (*V. dahuricum, V. grandiflorum,* and *V. nanchuanense*) lacked the protein-coding gene *rps16*. Compared to other genera in Liliales, the loss of the *rps16* gene seems to be a specific feature of species in tribes Melanthieae and Chionographideae (Bodin et al., 2013; Do et al., 2013; Han et al., 2019). In addition, our analysis indicated that *ycf68* and *ycf15* do not encode proteins and may, therefore, have been pseudogenized, as suggested in previous studies (Raubeson et al., 2007; Do et al., 2013). An additional potential pseudogene was identified in the chloroplast genomes of four *Veratrum* species (*V. dahuricum, V. grandiflorum*, *V. nanchuanense*, and *V. oxysepalum)*; the *infA* gene in these genomes contained an internal stop codon. The *infA* gene is the most common gene loss in angiosperm chloroplast genomes and is widely believed to have been transferred from the plastid to nuclear genome (Cummings and Hershey, 1994; Millen et al., 2001).

Synonymous codon usage bias of protein-coding genes in the complete cp genome is known to be an essential evolutionary character within the genome that affects gene expression and organism evolution (Li et al., 2019). In this study, amino acids with an A/U-end codon were preferred in *Veratrum*, consistent with codon usage bias in other members of Melanthiaceae (Do et al., 2013; Han et al., 2019; Yang et al., 2019). One explanation for this bias is that dicotyledons have greater preference for the A/U-end codons. If so, this may have contributed to the molecular evolution of *Veratrum* species, especially by natural selection and mutation pressure (Sharp and Wen-Hsiung, 1986; Bulmer, 1991).

Population genetics has long relied on SSRs derived from the plastid region because of their high variability and stable reproducibility. The SSRs identified in this study are potentially useful in developing lineage-specific plastid-based SSR markers for studies examining the population genetics of *Veratrum*. In addition, large

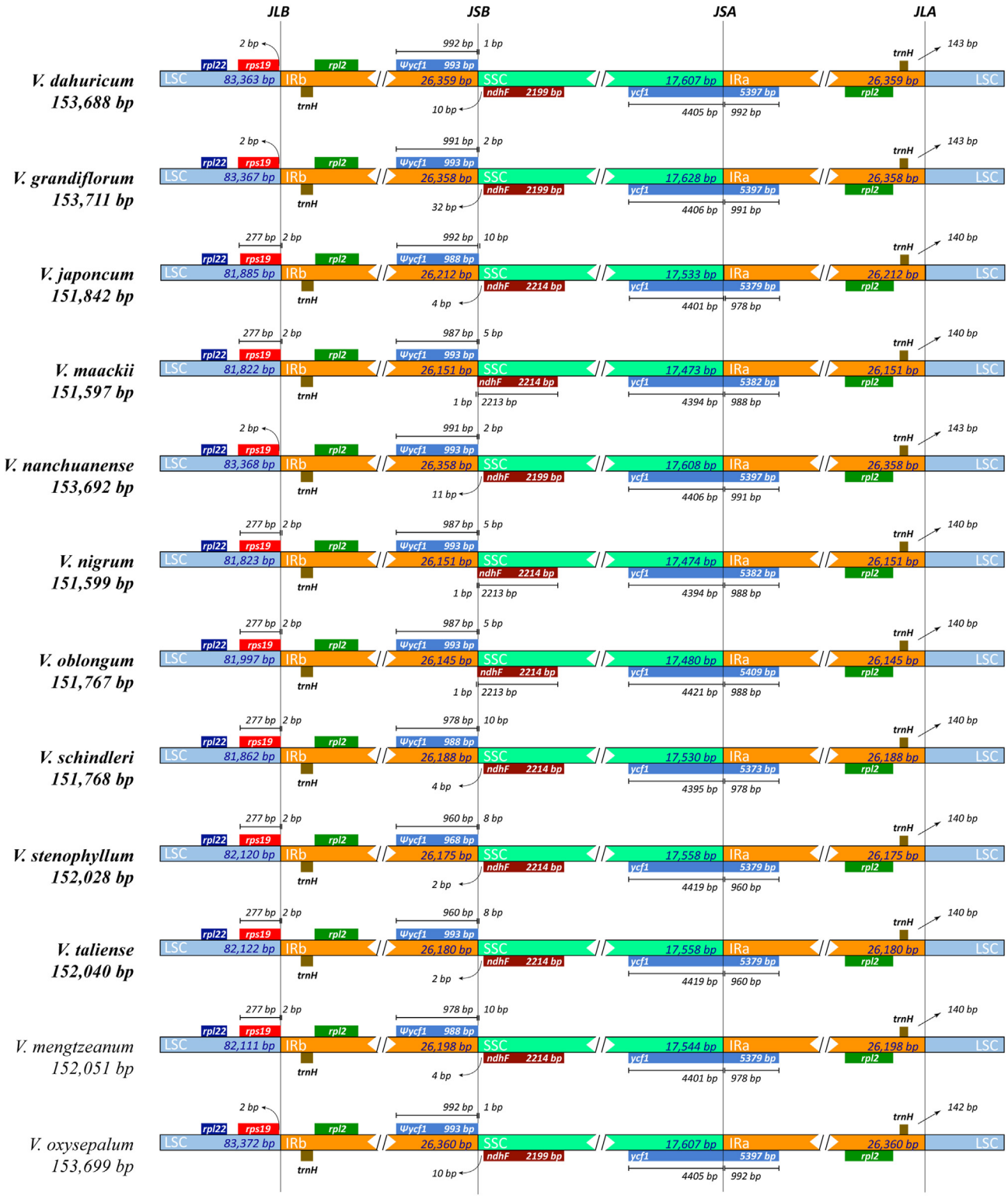


Fig. 8. Comparison of the LSC/IRb/SSC/IRa junctions among the twelve *Veratrum* cp genomes.

repeat sequences play an important role in the insertion/deletion mismatches and recombination that led to genomic variation (Asaf et al., 2018). Although not explored in this study, these large repeat motifs may provide useful information for the development of genetic markers that can be used to analyze population genetics (Nie et al., 2012).

4.2. Identification of highly variable regions

DNA barcodes and gene markers derived from the highly variable regions in the cp genome are effective tools that can be used to identify closely related species in the plant kingdom (Li et al., 2015). Previous studies have demonstrated that two fragments

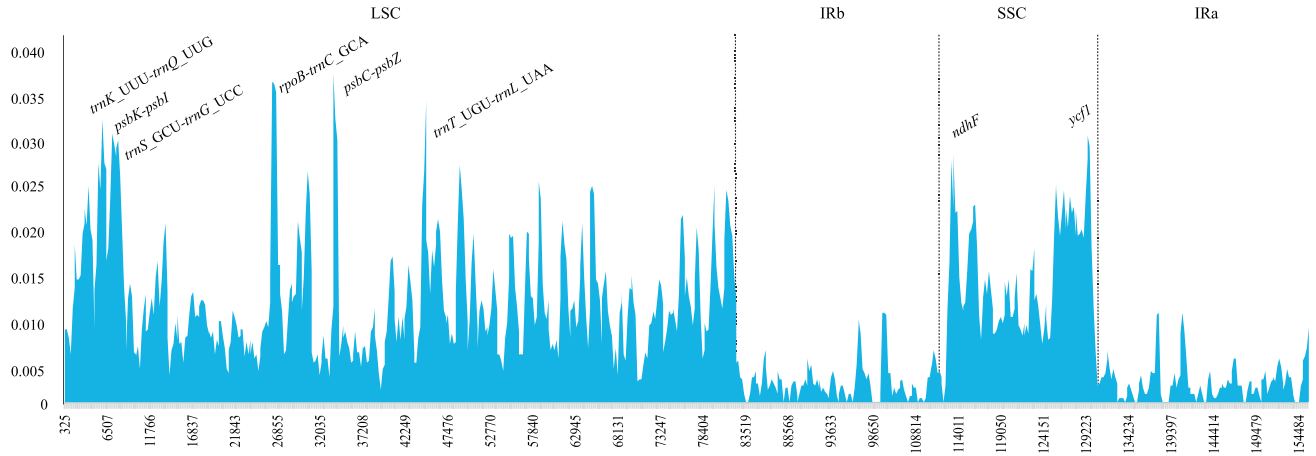


Fig. 9. Sliding-window analysis on the cp genomes for twelve *Veratrum* species. X-axis: position of the midpoint of a window; Y-axis: nucleotide diversity (Pi) of each window.

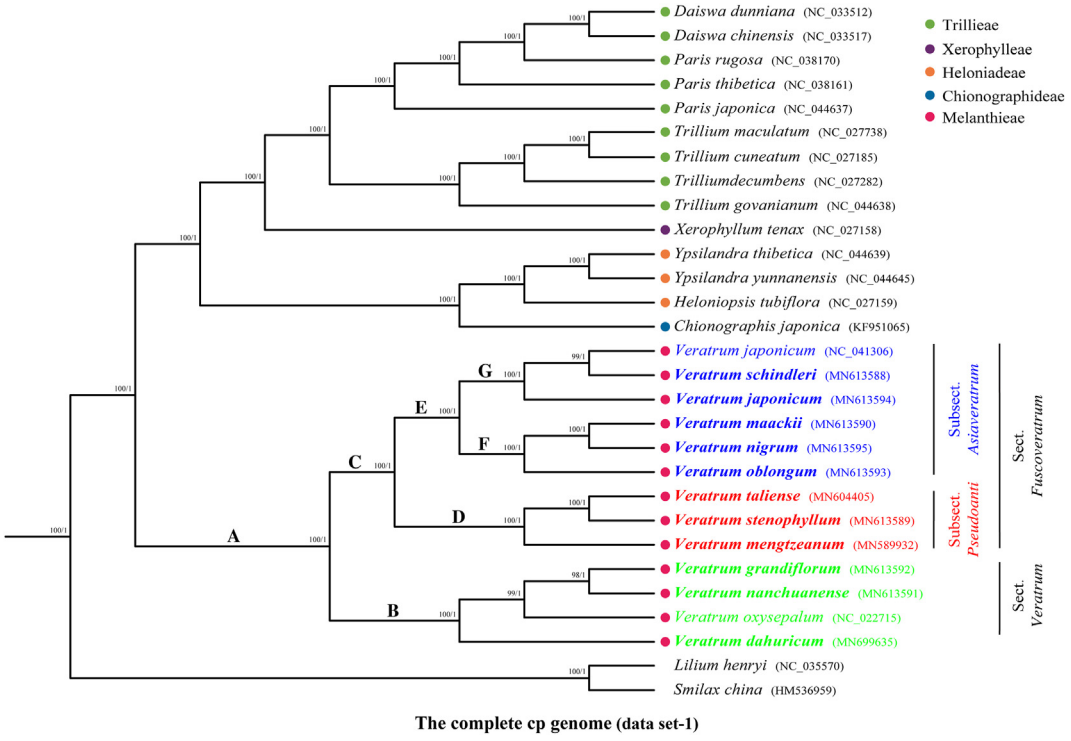


Fig. 10. Phylogenetic tree of 29 taxa using maximum likelihood (ML) and Bayesian inference (BI) methods based on cp genomes. ML tree topology is shown with ML bootstrap values, and BI posterior probabilities are indicated on the nodes. Colored lines and braces at the right of the tree indicate section and subgenus names of *Veratrum*. The color of the dot represents different tribes.

(*trnS-trnG* and *ndhF*) serve as useful markers for studies examining *Veratrum* phylogeny (Zomlefer et al., 2006; Liao et al., 2007; Kim et al., 2014, 2016). We identified at least eight highly variable regions in *Veratrum* cp genomes that may serve as powerful DNA barcodes for the genus, expanding the number of potential markers. However, future studies are required to determine which of these variable regions are suitable for species delimitation and identification.

4.3. Phylogenetic inferences

Previous studies failed to clearly elucidate the phylogenetic relationships in *Veratrum*, because they lacked sufficient polymorphic sites (Zomlefer et al., 2003; Kim et al., 2014, 2016). This limitation can be addressed by using complete cp genome sequences, which contain sufficient informative sites and can successfully resolve phylogenetic relationships at almost any taxonomic level (Jansen et al., 2006; Zhang et al., 2016). In this study, we used four sequence data sets to construct phylogenetic trees for *Veratrum*. Our classification of *Veratrum* at subgenus level are consistent using all four data sets (Figs. 10 and 11). Furthermore, our *Veratrum* phylogeny supports those of previous studies (Zomlefer et al., 2003; Kim et al., 2016). The APG II previously revised *Veratrum* at the subgenus level (Bremer et al., 2003), which has since been supported by studies that used short DNA markers (Zomlefer et al., 2006; Kim et al., 2016). Although without sufficient sampling

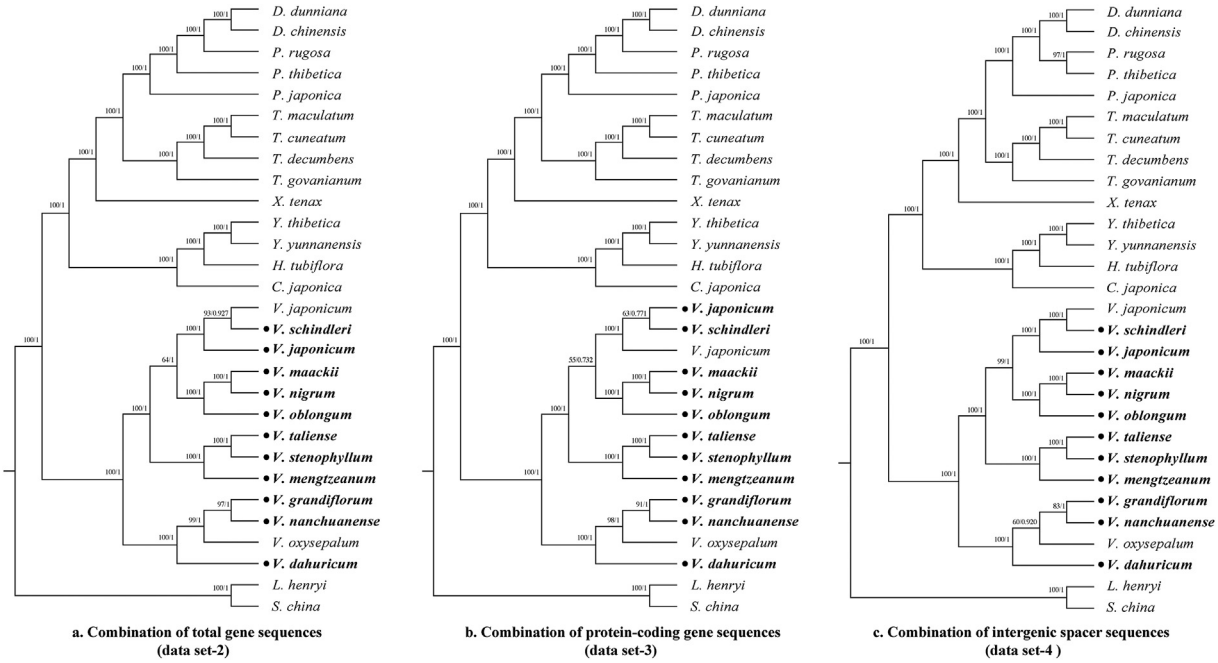


Fig. 11. Phylogenetic tree of 29 taxa using maximum likelihood (ML) and Bayesian inference (BI) methods based on the three data sets. (a) Combination of total gene sequences, including protein-coding, tRNA and rRNA genes (data set-2). (b) Combination of protein-coding gene sequences (data set-3). (c) Combination of intergenic region sequences (data set-4). ML tree topology is shown with ML bootstrap values, and BI posterior probabilities are indicated on the nodes.

size, the fact that *Veratrum* is monophyletic, based on its current circumscription, cannot be overlooked.

Our phylogenetic analysis clarified the position of *Veratrum nanchuanense* within *Veratrum* (Fig. 10, clade B). *Veratrum* species are separated into two different sections based on the leaf position and the type of style on the fruiting carpels (Zomlefer et al., 2003; Kim et al., 2014). However, these morphological features cannot be used to classify *V. nanchuanense*, which consequently holds an ambiguous systematic position (Tsi, 1980; Chen and Takahashi, 2000). The morphological characteristics (e.g., leafy stems and pubescent many-branched inflorescences) provide some support for including *V. nanchuanense* under the section *Veratrum*, our phylogenetic tree undoubtedly places *V. nanchuanense* in the section, sister to *V. grandiflorum* (Fig. 10).

Our phylogenetic analysis also highlights concerns about the safety of using certain species of *Veratrum* for medicinal purposes. Generally, the active ingredients of closely related species have similar chemical compositions and efficacies; however, the medicinal use of closely related species may have unintended adverse effects (Ma et al., 2019). Medical books have recorded the use of four *Veratrum* species in the production of a traditional folk medicine called "*pimacao*": *V. grandiflorum, V. mengtzeanum*, *V. stenophyllum*, and *V. taliense* (Wang et al., 2005). Here we demonstrate that one of these species is less closely related to the others than previously thought. Specifically, *V. grandiflorum* belongs to section *Veratrum*, whereas *V. stenophyllum*, *V. taliense* and *V. mengtzeanum* belong to section *Fuscoveratrum* (Fig. 10, clades B and C), raising concerns about the safety of using *V. grandiflorum* as a raw material in "*pimacao*". We recommend that future studies assess the chemical composition and pharmacological effects of *V. grandiflorum*.

Our study also resolved phylogenetic relationships within the *Veratrum* subsection *Asiaveratrum*. Morphological synapomorphies and the lack of high-resolution molecular markers have limited previous taxonomic treatments of the subsection (Satake 1982; Zomlefer et al., 2003; Liao et al., 2007). For example, in previous studies, all species of subsection *Asiaveratrum* formed an unresolved polytomy among *V. maackii*, *V. nigrum*, *V. japonicum*, and *V. schindleri* (Zomlefer et al., 2003). In this study, we detected high resolution relationships for all taxa within subsection *Asiaveratrum*. Specifically, we found that *V. maackii* and *V. nigrum* are closely related, and form a sister clade of *V. oblongum* with strong support (Fig. 10). Morphologically, *V. nigrum*, which has sessile leaves or sometimes shortly petiolate in the distal part of the stem, differs clearly from *V. maackii* (Chen and Takahashi, 2000). Hence, we suggest treating *V. nigrum* as a species distinct from *V. maackii* (Fig. 10, clade G). In addition, the taxonomic status of *V. japonicum* has long been controversial. *V. japonicum* specimens from Yunnan Province coincide with descriptions of this species in *Flora Reipublicae Popularis Sinicae*; however, specimens from Fujian Province coincide with descriptions of *V. schindleri* (Tsi, 1980). Our molecular phylogenetic analysis indicates that *V. japonicum* (Yunnan, China) is nested within one clade comprising *V. japonicum* (Korea) and *V. schindleri*. These finding show that *V. japonicum* – both the collections designated as *V. schindleri* in Flora of China (Chen and Takahashi, 2000) and the collection from Korea – should be considered as conspecific with *V. schindleri*. Thus, we recommend reducing *V. japonicum* to *V. schindleri*.

5. Conclusions

In this study, we investigated ten complete cp genomes of *Veratrum* species from East Asia, representing the majority of species distributed in China and about one-quarter of the global species of *Veratrum*. The cp genomes exhibited a typical circular quadripartite structure and had highly similar gene order and genomic structure. SSRs, long repeats, and the eight highly variable loci were identified across the *Veratrum* cp genomes, which could serve as potential markers for phylogenetic and population genetics studies. The monophyly of the genus *Veratrum* was confirmed and phylogenetic analysis indicated that the genus consists of two sections. The circumscription and relationships of

infrageneric taxa of *Veratrum* were well evaluated, too. These results provide a better understanding of the phylogenetic relationships in the genus *Veratrum* and afford more genomic information for further utilization of medicinal species in the genus *Veratrum* and for the evolutionary studies of Melanthiaceae.

Author contributions

Conceptualization, G.L., Z.Q.; software, Y.Z., L.H., and C.Y.; formal analysis, Y.Z., L.H., and X.T.; investigation, G.L.,Z.Q., C.Y. and Z.Y.; resources, G.L., C.Y., Z.Y. and Y.Z.; data curation, L.H., Z.Y., and X.T.; writing—original draft preparation, Y.Z., L.H., and C.Y. All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

We are very grateful to Dr. Yong-Hong Zhang of Yunnan Normal University for his comments and technical supports. We thank Dr. Shiou Yih Lee of Sun Yat-sen University for ideas and assistance provided during manuscript preparation. This work was funded by the High-end talent of Yunnan province 'Youth Top-notch Talent Support Program' the Reserve Talent for Academic and Young Technical Leaders of Yunnan Province Special Subsidies for Public Health services of "Fourth Chinese Materia Medica Resources Survey" (DSS, MOF. 43/2018).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pld.2021.05.004.

References

Aii, J., Kishima, Y., Mikami, T., et al., 1997. Expansion of the IR in the chloroplast genomes of buckwheat species is due to incorporation of an SSC sequence that could be mediated by an inversion. Curr. Genet. 31, 276–279. https://doi.org/10.1007/s002940050206.

Amiryousefi, A., Hyvonen, J., Poczai, P., 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics 34, 3030–3031. https://doi.org/10.1093/bioinformatics/bty220.

Asaf, S., Khan, A.L., Khan, M.A., et al., 2018. Complete chloroplast genome sequence and comparative analysis of loblolly pine (*Pinus taeda* L.) with related species. PLoS One 13, e0192966. https://doi.org/10.1371/journal.pone.0192966.

Bodin, S.S., Kim, J.S., Kim, J.H., 2013. Complete chloroplast genome of *Chionographis japonica* (Willd.) Maxim. (Melanthiaceae): comparative genomics and evaluation of universal primers for Liliales. Plant Mol. Biol. Rep. 31, 1407–1421. https://doi.org/10.1007/s11105-013-0616-x.

Bodkin, N.L., 1978. A Revision of North American *Melanthium* L. (Liliaceae). Ph.D. dissertation. University of Maryland, College Park, Maryland, USA.

Bremer, B., Bremer, K., Chase, M.W., et al., 2003. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141, 399–436. https://doi.org/10.1046/j.1095-8339.2003.t01-1-00158.x.

Bulmer, M.G., 1991. The selection-mutation-drift theory of synonymous codon usage. Genetics 129, 897–907. https://doi.org/10.1016/1050-3862(91)90016-K.

Chandler, C.M., McDougal, O.M., 2014. Medicinal history of north American *Veratrum*. Phytochemistry Rev. 13, 671–694. https://doi.org/10.1007/s11101-013-9328-y.

Chase, M.W., Christenhusz, M.J.M., Fay, M.F., et al., 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181, 1657–1669. https://doi.org/10.1111/boj.12385.

Chen, X.Q., Takahashi, H., 2000. *Veratrum* Linnaeus. In: Wu, Z.Y., Raven, P.H., Hong, D.Y. (Eds.), Flora of China, vol. 24. Science Press, Beijing, China, pp. 82–85.

Chumley, T.W., Palmer, J.D., Mower, J.P., et al., 2006. The complete chloroplast genome sequence of *Pelargonium* × *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol. Biol. Evol. 11, 2175–2190. https://doi.org/10.1093/molbev/msl089.

Cong, Y., Wu, Y., Shen, S., et al., 2020. A structure-activity relationship between the *Veratrum* alkaloids on the antihypertension and DNA damage activity in mice. Chem. Biodivers. 17, e1900473 https://doi.org/10.1002/cbdv.201900473.

Cummings, H.S., Hershey, J.W.B., 1994. Translation initiation factor IF1 is essential for cell viability in *Escherichia coli*. J. Bacteriol. 176, 198–205. https://doi.org/10.1128/jb.176.1.198-205.1994.

Dierckxsens, N., Mardulyn, P., Smits, G., 2017. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45, e18. https://doi.org/10.1093/nar/gkw955.

Do, H.D.K., Kim, J.S., Kim, J.H., 2013. Comparative genomics of four Liliales families inferred from the complete chloroplast genome sequence of *Veratrum patulum* O. Loes. (Melanthiaceae). Gene 530, 229–235. https://doi.org/10.1016/j.gene.2013.07.100.

Doyle, J.J., Davis, J.I., Soreng, R.J., et al., 1992. Chloroplast DNA inversions and the origin of the grass family (Poaceae). Proc. Natl. Acad. Sci. U.S.A. 89, 7722–7726. https://doi.org/10.1073/pnas.89.16.7722.

Frazer, K.A., Pachter, L., Poliakov, A., et al., 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32, W273–W279. https://doi.org/10.1093/nar/gkh458.

Greiner, S., Lehwark, P., Bock, R., 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47, W59–W64. https://doi.org/10.1093/nar/gkz238.

Gu, C., Ma, L., Wu, Z., et al., 2019. Comparative analysis of chloroplast genomes from 22 Lythraceae species: inferences for phylogenetic relationships and genome evolution within Myrtales. BMC Plant Biol. 19, 281. https://doi.org/10.1186/s12870-019-1870-3.

Han, L.J., Liu, Y.Y., Zhang, Y.M., et al., 2019. The complete chloroplast genome and phylogenetic analysis of *Veratrum mengtzeanum* Loes. F. (Liliaceae). Mitochondrial DNA 4, 4170–4171. https://doi.org/10.1080/23802359.2019.1693926.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754.

Jansen, R.K., Kaittanis, C., Saski, C., et al., 2006. Phylogenetic analysis of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. BMC Evol. Biol. 6, 32. https://doi.org/10.1186/1471-2148-6-32.

Kalyaanamoorthy, S., Minh, B., Wong, T., et al., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Kearse, M., Moir, R., Wilson, A., et al., 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Kikkawa, H.S., Tsuge, K., Kubota, S., et al., 2017. Species identification of white false hellebore (*Veratrum album* subsp. *oxysepalum*) using real-time PCR. Forensic Sci. Int. 275, 160–166. https://doi.org/10.1016/j.forsciint.2017.02.002.

Kim, K.J., Lee, H.L., 2004. Complete chloroplast genome sequence from Korean Ginseng (Panax schiseng Nees) and comparative analysis of sequence evolution among 17 vascular plants. DNA Res 11, 247–261. https://doi.org/10.1093/dnares/11.4.247.

Kim, J.O., Tamura, M.N., Fuse, S., et al., 2014. Taxonomic status and phylogeny of *Veratrum* section *Veratrum* (Melanthiaceae) in Korea and Japan based on chloroplast and nuclear sequence data. Plant Systemat. Evol. 300, 75–89. https://doi.org/10.1007/s00606-013-0861-3.

Kim, S.C., Kim, J.S., Chase, M.W., et al., 2016. Molecular phylogenetic relationships of Melanthiaceae (Liliales) based on plastid DNA sequences. Bot. J. Linn. Soc. 181, 567–584. https://doi.org/10.1111/boj.12405.

Kumar, S., Stecher, G., Li, M., et al., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549. https://doi.org/10.1093/molbev/msy096.

Kupchan, S.M., Zimmerman, J.H., Afonso, A., 1961. The alkaloids and taxonomy of *Veratrum* and related genera. Lloydia 24, 1–26.

Kurtz, S., Choudhuri, J.V., Ohlebusch, E., et al., 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 29 https://doi.org/10.1093/nar/29.22.4633, 4633–4622.

Kurtz, S., Phillippy, A., Delcher, A.L., et al., 2004. Versatile and open software for comparing large genomes. Genome Biol. 5, R12. https://doi.org/10.1186/gb-2004-5-2-r12.

Li, X., Yang, Y., Henry, R.J., et al., 2015. Plant DNA barcoding: from gene to genome. Biol. Rev. Camb. Phil. Soc. 90, 157–166. https://doi.org/10.1111/brv.12104.

Li, Q., Yang, K.X., Zhao, Y.L., et al., 2016. Potent anti-inflammatory and analgesic steroidal alkaloids from *Veratrum taliense*. J. Ethnopharmacol. 179, 274–279. https://doi.org/10.1016/j.jep.2015.12.059.

Li, Y., Sylvester, S.P., Li, M., et al., 2019. The complete plastid genome of *Magnolia zenii* and genetic comparison to Magnoliaceae species. Molecules 24, 261. https://doi.org/10.3390/molecules24020261.

Liao, W.J., Yuan, Y.M., Zhang, D.Y., 2007. Biogeography and evolution of flower color in *Veratrum* (Melanthiaceae) through inference of a phylogeny based on multiple DNA markers. Plant Systemat. Evol. 267, 177–190. https://doi.org/10.1007/s00606-007-0528-z.

Linnaeus, C., 1753. Species plantarum. Stockholm, Holmiae, Impensis L. Salvii. 1, 2.

Loesener, O., 1926. Studien über die Gattung *Veratrum* und ihre Verbreitung. Verh. Bot. Ver. Prov. Brandenb. 68, 105–166.
Loesener, O., 1927. Übersicht über Arten der Gattung *Veratrum*, Teil I. Feddes Repert. Specierum Nov. Regni Veg. 24, 61–72.
Loesener, O., 1928. Übersicht über Arten der Gattung *Veratrum*, Schluss. Feddes Repert. Specierum Nov. Regni Veg. 25, 1–10.
Ma, H.F., Sima, Y.K., Zhang, D., et al., 2019. Chemical constituents of the volatile oils from the leaves of three *Michelia* species. J. Northwest For. Univ. 34, 212–216.
Millen, R.S., Olmstead, R.G., Adams, K.L., et al., 2001. Many parallel losses of infa from chloroplast dna during angiosperm evolution with multiple independent transfers to the nucleus. Plant Cell 13, 645–658. https://doi.org/10.1105/tpc.13.3.645.
Moore, M.J., Soltis, P.S., Bell, C.D., et al., 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc. Natl. Acad. Sci. U.S.A. 107, 4623–4628. https://doi.org/10.1073/pnas.0907801107.
Mudunuri, S.B., Nagarajaram, H.A., 2007. IMEx: imperfect microsatellite extractor. Bioinformatics 23, 1181–1187. https://doi.org/10.1093/bioinformatics/btm097.
Nakai, T., 1937a. Species generic Veratri in regione Manshurico-Koreano sponte nascentes. Rep. Inst. Sci. Res. Manchoukuo 325–344.
Nakai, T., 1937b. Japanese species of *Veratrum*. Shokubutsu Kenkyu Zasshi 13 (631–645), 701–713.
Nie, X., Lv, S., Zhang, Y., et al., 2012. Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*). PLoS One 7, 36869. https://doi.org/10.1371/journal.pone.0036869.
Patel, R.K., Jain, M., 2012. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7, e30619. https://doi.org/10.1371/journal.pone.0030619.
Raubeson, L.A., Peery, R., Chumley, T.W., et al., 2007. Comparative chloroplast genomics: analysis including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. BMC Genom. 8, 174. https://doi.org/10.1186/1471-2164-8-174.
Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., et al., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302. https://doi.org/10.1093/molbev/msx248.
Satake, Y., 1982. *Veratrum*. In: Satake, Y., Ohwi, J., Kitamura, S., et al. (Eds.), Wild Flowers of Japan, Herbaceous Plants. Heibonsha, Tokyo, pp. 28–29.
Sharp, P.M., Wen-Hsiung, L., 1986. Codon usage in regulatory genes in, *Escherichia coli* does not reflect selection for 'rare' codons. Nucleic Acids Res. 19, 7737–7749. https://doi.org/10.1093/nar/14.19.7737.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
Sugiura, M., 1992. The chloroplast genome. Plant Mol. Biol. 19, 149–157.
Tamura, M.N., 1998. Melanthiaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants. Flowering Plants: Monocotyledons. Springer, Berlin, pp. 369–380.
Tsi, Z.H., 1980. Veratrum. In: Flora Reipublicae Popularis Sinicae (Ed.), Delectis Florae Reipublicae Popularis Sinicae Agenda Academiae Sinicae Edita, vol. 14. Science Press, Beijing, China, pp. 19–30.
Turner, M.W., Rossi, M., Campfield, V., et al., 2019. Steroidal alkaloid variation in *Veratrum californicum* as determined by modern methods of analytical analysis. Fitoterapia 137, 104281. https://doi.org/10.1016/j.fitote.2019.104281.
Wang, Q.X., Wang, R.S., Yue, K.L., et al., 2005. Standard of Chinese Herbal Pieces in Yunnan Province. Yunnan Art Publishing Press, Kunming, China, pp. 50–51.
Wicke, S., Schneeweiss, G.M., De Pamphilis, C.W., et al., 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol. Biol. 76, 273–297. https://doi.org/10.1007/s11103-011-9762-4.
Wolf, P.G., Roper, J.M., Duffy, A.M., 2010. The evolution of chloroplast genome structure in ferns. Genome 53, 731–738. https://doi.org/10.1139/g10-061.
Wu, Y.R., Song, L.R., Hu, L., et al., 1999. Chinese Materia Medica (Zhonghua Bencao), vol. 8. Shanghai Science and Technology Press, Shanghai, China, p. 183.
Wyman, S.K., Jansen, R.K., Boore, J.L., 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20, 3252–3255. https://doi.org/10.1093/bioinformatics/bth352.
Xie, D.F., Yu, H.X., Price, M., et al., 2019. Phylogeny of Chinese *Allium* species in section *Daghestanica* and adaptive evolution of *Allium* (Amaryllidaceae, Allioideae) species revealed by the chloroplast complete genome. Front. Plant Sci. 10, 460. https://doi.org/10.3389/fpls.2019.00460.
Yang, J., Yang, L., Yang, Z., et al., 2019. Chloroplast phylogenomic analysis provides insights into the evolution of the largest eukaryotic genome holder, *Paris japonica* (Melanthiaceae). BMC Plant Biol. 19, 293. https://doi.org/10.1186/s12870-019-1879-7.
Yang, Z., Zhao, T., Ma, Q., et al., 2018. Comparative genomics and phylogenetic analysis revealed the chloroplast genome variation and interspecific relationships of *Corylus* (Betulaceae) species. Front. Plant Sci. 9, 927. https://doi.org/10.3389/fpls.2018.00927.
Yin, Z.L., Xie, H., Zhang, J., 2014. A preliminary study of biological characteristics of *Veratrum nigrum*. Yunnan J. Tradit. Chin. Med. Mater. Med. 4, 54–57. https://doi.org/10.16254/j.cnki.53-1120/r.2014.03.035.
Zhang, Y., Du, L., Ao, L., et al., 2016. The complete chloroplast genome sequences of five *Epimedium* species: lights into phylogenetic and taxonomic analysis. Front. Plant Sci. 7, 306. https://doi.org/10.3389/fpls.2016.00306.
Zimmerman, J.H., 1958. A Monograph of *Veratrum*. University of Wisconsin, Madison, Wisconsin, USA. Ph. D. thesis.
Zomlefer, W.B., 1997. The genera of Melanthiaceae in the southeastern United States. Harv. Pap. Bot. 2, 133–177. http://www.jstor.org/stable/41761544.
Zomlefer, W.B., Williams, N.H., Whitten, W.M., et al., 2001. Generic circumscription and relationships in the tribe Melanthieae (Liliales, Melanthiaceae), with emphasis on *Zigadenus* : evidence from ITS and *trnL-F* sequence data. Am. J. Bot. 88, 1657–1669. https://doi.org/10.2307/3558411.
Zomlefer, W.B., Whitten, W.M., Williams, N.H., et al., 2003. An overview of *Veratrum* s.l. (Liliales: Melanthiaceae) and an infrageneric phylogeny based on ITS sequence data. Syst. Bot. 28, 250–269. https://doi.org/10.1043/0363-6445-28.2.250.
Zomlefer, W.B., Judd, W., Whitten, M., et al., 2006. A synopsis of Melanthiaceae (Liliales) with focus on character evolution in tribe Melanthieae. Aliso 22, 566–578. https://doi.org/10.5642/aliso.20062201.44.