Title

Molecular phylogeny and inflorescence evolution of Prunus (Rosaceae) based on RAD-seq and genome skimming analyses.

Journal

Plant Diversity, 45(4)

Authors

Su, Na
Hodel, Richard
Wang, Xi
et al.

Publication Date

2023-07-01

DOI

10.1016/j.pld.2023.03.013

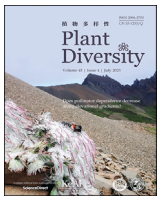
Peer reviewed

Plant Diversity 45 (2023) 397–408

Contents lists available at ScienceDirect

Plant Diversity

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



Research paper

Molecular phylogeny and inflorescence evolution of *Prunus* (Rosaceae) based on RAD-seq and genome skimming analyses

Na Su [a, b, 1], Richard G.J. Hodel [c, 1], Xi Wang [a, b, 1], Jun-Ru Wang [a, b, 1], Si-Yu Xie [a, b], Chao-Xia Gui [a, b], Ling Zhang [d], Zhao-Yang Chang [a, b], Liang Zhao [a, b, *], Daniel Potter [e], Jun Wen [c]

[a] College of Life Sciences, Northwest A&F University, Yangling 712100, China
[b] Herbarium of Northwest A&F University, Yangling 712100, China
[c] Department of Botany, National Museum of Natural History, MRC 166, Smithsonian Institution, Washington, DC 20013-7012, USA
[d] College of Life Sciences, Tarim University, Alaer 843300, China
[e] Department of Plant Sciences, MS2, University of California, Davis, CA 95616, USA

ARTICLE INFO

Article history:
Received 6 September 2022
Received in revised form
29 March 2023
Accepted 31 March 2023
Available online 6 April 2023

Keywords:
Prunus
Rosaceae
RAD-Seq
Chloroplast genome
Hybridization
Inflorescence evolution

ABSTRACT

Prunus is an economically important genus widely distributed in the temperate Northern Hemisphere. Previous studies on the genus using a variety of loci yielded conflicting phylogenetic hypotheses. Here, we generated nuclear reduced representation sequencing data and plastid genomes for 36 *Prunus* individuals and two outgroups. Both nuclear and plastome data recovered a well-resolved phylogeny. The species were divided into three main clades corresponding to their inflorescence types, - the racemose group, the solitary-flower group and the corymbose group - with the latter two sister to one another. *Prunus* was inferred to have diversified initially in the Late Cretaceous around 67.32 million years ago. The diversification of the three major clades began between the Paleocene and Miocene, suggesting that paleoclimatic events were an important driving force for *Prunus* diversification. Ancestral state reconstructions revealed that the most recent common ancestor of *Prunus* had racemose inflorescences, and the solitary-flower and corymb inflorescence types were derived by reduction of flower number and suppression of the rachis, respectively. We also tested the hybrid origin hypothesis of the racemose group proposed in previous studies. *Prunus* has undergone extensive hybridization events, although it is difficult to identify conclusively specific instances of hybridization when using SNP data, especially deep in the phylogeny. Our study provides well-resolved nuclear and plastid phylogenies of *Prunus*, reveals substantial cytonuclear discord at shallow scales, and sheds new light on inflorescence evolution in this economically important lineage.

1. Introduction

Angiosperms exhibit a great diversity of inflorescence types, such as racemes, panicles, corymbs, and cymes, depending on the number and arrangement of flowers on shoots, and their branching patterns (Benlloch et al., 2007; Endress, 2010). The evolution of inflorescences has attracted the interest of botanists for a long time (e.g., Parkin, 1914; Stebbins, 1973; Takhtajan, 1991; Endress, 2010). The rapid development of DNA sequencing technology, computational resources, and phylogenetic inference methods provide an opportunity for inferring accurate phylogenies and conducting rigorous tests of hypotheses concerning inflorescence evolution for some angiosperm groups based on genomic and phylogenetic data (Endress, 2010; Gerrath et al., 2017; Ma et al., 2017).

Prunus L. consists of 250–400 species of trees and shrubs widely distributed in the northern temperate zone and subtropical and tropical regions (Rehder, 1956; Yü et al., 1986; Lu et al., 2003), with eastern Asia being its center of diversity (Wen et al., 2008; Chin et al.,

* Corresponding author. College of Life Sciences, Northwest A&F University, Yangling, 712100, China.
E-mail address: biology_zhaoliang@126.com (L. Zhao).
Peer review under responsibility of Editorial Office of Plant Diversity.
[1] Na Su, Richard G. J. Hodel, Xi Wang and Junru Wang contributed equally to this work.

https://doi.org/10.1016/j.pld.2023.03.013

2014). This genus is characterized by simple leaves with stipules and leaf glands, a superior single ovary, and drupe fruits (Lu et al., 2003; Zhao et al., 2016). Many species of *Prunus* are economically significant as food crops, such as peach, plum, almond, and sweet cherry (Fig. 1; Bortiri et al., 2001; Lee and Wen, 2001), and many others have been used as ornamentals, timber, and medicine (Andro and Riffaud, 1995; Lee and Wen, 2001; Wen et al., 2008). Species of *Prunus* show a diversity of floral displays with several distinct inflorescence types, including racemes, corymbs, and solitary flowers, making it an ideal group to investigate the evolutionary transitions among different inflorescence organizations.

Rehder (1956) divided *Prunus* into five subgenera, namely *Amygdalus* L., *Cerasus* Mill., *Lauro-cerasus* Tourn. ex Duhamel, *Padus* Mill. and *Prunus*, and several taxonomists have accepted this classification (Kalkman, 2004; Wen et al., 2008; Chin et al., 2010, 2014; Zhao et al., 2016). Over the last two decades, researchers have tried to reconstruct the phylogenetic relationships of *Prunus* using different genomic regions, such as plastid markers, nuclear ribosomal ITS, and other nuclear loci (Bortiri et al., 2001, 2002, 2006; Lee and Wen, 2001; Wen et al., 2008; Chin et al., 2010, 2014; Shi et al., 2013; Zhao et al., 2016, 2018; Hodel et al., 2021). These analyses have recovered three main groups within *Prunus*, each of

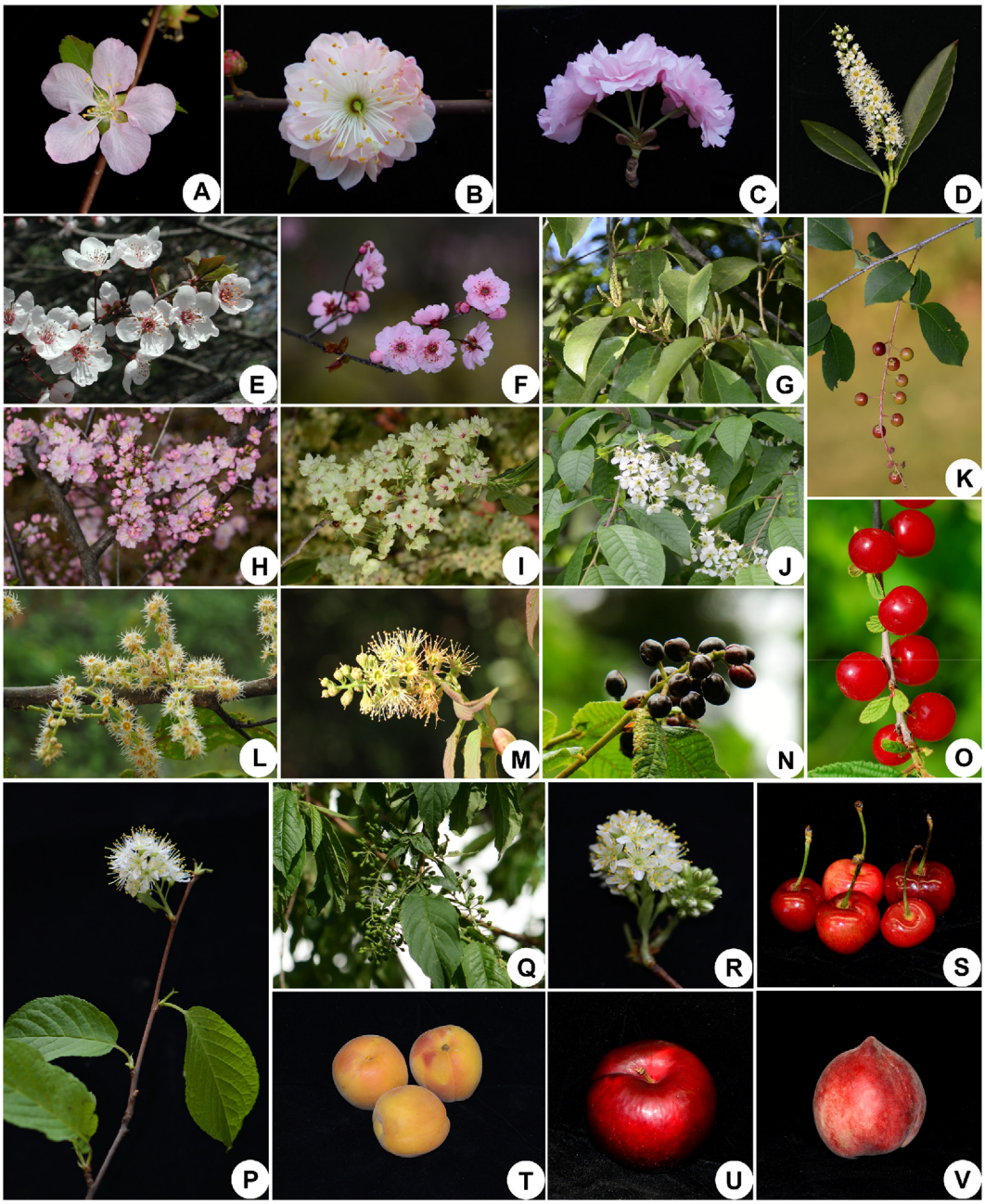


Fig. 1. Morphological diversity of *Prunus* species. (A) *P. triloba*; (B) *P. triloba* f. *multiplex*; (C) *P. serrulata*; (D) *P. laurocerasus*; (E) *P. cerasifera* f. *atropurpurea*; (F) *P. mume*; (G) *P. zippeliana*; (H) *P. glanulosa*; (I) *P. serrulata* 'Grandiflora'; (J) *P. padus*; (K) *P. serotina*; (L) *P. henryi*; (M–N) *P. hypoxantha*; (O) *P. tomentosa*; (P–R) *P. maackii*; (S) cherry; (T) apricot; (U) plum; (V) peach.

them presenting a synapomorphic inflorescence structure: (1) the solitary-flower group consisting of *Prunus* subg. *Amygdalus* (L.) Foche and *P.* subg. *Prunus* (including section *Armeniaca* (Scop.) Nakai); (2) the corymbose group, which includes *P.* subg. *Cerasus* (Mill.) Pers; and (3) the racemose group comprising *P.* subg. *Laurocerasus* and *P.* subg. *Padus*, as well as the *Maddenia* Hook. f. & Thomson and *Pygeum* Gaertn. groups (Chin et al., 2014). Most species of the solitary-flower and corymbose groups are diploid, whereas taxa of the racemose group usually have higher ploidy levels (Chin et al., 2014; Zhao et al., 2016; Hodel et al., 2021; CCDB, http://ccdb.tau.ac.il/). Nevertheless, the frequent, ancient hybridization events in the evolutionary history of *Prunus* have greatly challenged the clarification of the phylogenetic relationships among these three groups. The monophyly of the racemose group has been supported by plastid sequences (e.g., Chin et al., 2010, 2014); however, most analyses of nuclear sequence data have resolved the racemose group as paraphyletic (e.g., Bortiri et al., 2001, 2002; Lee and Wen, 2001; Chin et al., 2014), leading to hypotheses of multiple hybrid origins of this mostly polyploid lineage (Chin et al., 2014; Zhao et al., 2016). The monophyly of the racemose group has been strongly supported by hundreds of single-copy nuclear genes and chloroplast genomes from 21 transcriptomic data points (Hodel et al., 2021); however, only two species of the racemose group were included in their study, questioning the representativeness of the sampling regarding the taxonomic and morphological diversity in this lineage. Therefore, these results need to be further tested using more comprehensive taxon sampling.

Due to limited taxon sampling and informative loci in previous molecular phylogenetic studies of *Prunus*, the phylogenetic relationships among and within the major lineages still need to be reevaluated. Although phylogenomic data from transcriptomes and plastomes have generated a well-supported backbone for Rosaceae, limited taxon sampling of *Prunus* in these analyses failed to untangle all relationships (Xiang et al., 2016; Zhang et al., 2017). Moreover, the divergence times of some lineages within *Prunus* estimated in previous studies need to be further tested. Thus, exploring phylogenetic relationships within *Prunus* with extensive taxon sampling and molecular characters is essential.

The plastid genome (plastome) is characterized by its conserved structure and typically maternal inheritance, which makes it helpful in reconstructing phylogenetic relationships at both shallow and deep taxonomic levels (e.g., Li et al., 2019, 2021a; Liu et al., 2019, 2020a, 2020b; Li et al., 2020; Thode et al., 2020; Walker et al., 2022; Wang et al., 2020; Bai et al., 2021; Lei et al., 2021; Li et al., 2021b; Wu et al., 2021; Su et al., 2021; Lee et al., 2022; Liu et al., 2022a, 2022b; Xu et al., 2022; Zhang et al., 2022). However, the uniparental inheritance of plastomes limits their power to fully elucidate the evolutionary histories of lineages with pervasive reticulate evolution, such as the family Rosaceae. Therefore, nuclear and plastome data are needed to infer robust phylogenetic reconstruction. Restriction-site associated DNA sequencing (RAD-seq) can readily generate thousands of unlinked nuclear loci that can resolve dense reticulation on shallow systematic levels (e.g., Vargas et al., 2017; Ma et al., 2018; Mu et al., 2020; Zhou et al., 2020; Hodel et al., 2022).

In this study, we used plastomes from genome skimming data and single nucleotide polymorphism (SNPs) data from RAD-seq data to infer the phylogeny of *Prunus*, and we also reconstructed the evolution of inflorescence types in *Prunus*. We aim to (1) resolve phylogenetic relationships among the major lineages, (2) detect potential hybridization events, (3) estimate divergence times, and (4) infer the ancestral states and evolutionary transitions among the states of inflorescence types in *Prunus*.

2. Materials and methods

2.1. Sampling, DNA extraction, library preparation and sequencing

We sampled 38 accessions, including 32 *Prunus* species and two outgroup species (Table 1), for specific-locus amplified fragment sequencing (SLAF-seq), which is a reduced-representation genomic sequencing approach. Our sampling was taxonomically and morphologically representative, as these 32 *Prunus* species represented all subgenera and inflorescence types of *Prunus*. These species are mainly from Asia, which is considered the origin and diversity center of *Prunus* (Chin et al., 2014). SLAF-seq is a version of RAD-seq that uses a pre-determined reduced representation scheme to optimize and evenly space loci (Sun et al., 2013). For the plastome data, our sampling was identical to the sampling used for nuclear data, with 37 accessions newly sequenced for this study and one (*Prunus davidiana* (Carrière) Franch., accession number: MH460864) downloaded from GenBank. Total genomic DNA was extracted from the silica-gel-dried leaves using the CTAB method (Doyle and Doyle, 1987).

For RAD-seq data, the *Prunus mume* (Siebold) Siebold & Zucc. genome (accession number: GCF_000346735.1) was used in an *in silico* double enzyme digestion to determine the enzyme pair before preparing the library. The library preparation of RAD-seq was completed with the *Rsa*I + *Hae*III enzyme pair. The paired-end (264–414 bp) sequencing was carried out on the Illumina Hiseq™ 2500 sequencing platform (Illumina, Inc; San Diego, CA, USA) at Beijing Biomarker Technologies Corporation (Beijing, China). For the plastome data, the libraries were prepared in the Molecular Biology Experiment Center, Germplasm Bank of Wild Species in Southwest China using NEBNext® Ultra™ II DNA Library Prep Kit. Paried-end sequencing was conducted on a HiSeq 2500™ (Illumina) in BGI (Shenzhen, China).

2.2. SNP calling and plastome assembly

Raw sequencing reads were assembled *de novo* using ipyrad v.0.9.74 (Eaton, 2014; Eaton and Overcast, 2020) with default parameters, except the clustering threshold set at 0.85 and the maximum alleles per site set at 4. The resulting assembly consisted of 25,770,082 nucleotides per species representing 100,579 loci. This unfiltered 100,579 locus alignment was used in subsequent coalescent- and concatenation-based phylogenetic inferences. Additionally, the ipyrad assembly was filtered using vcftools (Danacek et al., 2011) to remove loci with minor allele frequency less than 0.05 (–maf = 0.05), and using several filtering levels for missing data to investigate the impact of missing data on the downstream analyses. We used three filtering levels, allowing a maximum of 40%, 60%, or 80% missing data per site (–max-missing = 0.6, 0.4, 0.2). These three SNP matrices were also used in the downstream phylogenetic analyses. We used BLAST searches to filter the RAD-seq data to remove any chloroplast- and mitochondrion-related loci with references from NCBI GenBank (*P. mume* chloroplast genome (accession number: NC_023798.1) and mitochondrial genome (accession number: NC_060491.1)).

Raw reads generated by genome skimming (Zhang et al., 2015; Liu et al., 2021) were assembled into plastomes using the GetOrganelle pipeline (Jin et al., 2020). The word sizes were set as 125 bp, round numbers were 15, and *k*-mers were 75, 85, 95 and 105. Complete plastid genomes from the following species were used as seed sequences for the assembly of all sequenced accessions: *Prunus kansuensis* Rehder (accession number: NC_023956), *Prunus maximowiczii* Rupr (accession number: NC_026981), *P. mume* (accession number: NC_023,798), *Prunus padus* L. (accession number: NC_026982), *Prunus persica* (L.) Batsch (accession

Table 1
Voucher information and GenBank accession numbers of sampled species of *Prunus* and outgroups.

Major group	Taxon	Voucher	Location	Inflorescence type	Latitude (N)	Longitude (E)	Altitude (m)	SRA accession number	
								RAD-seq	plastome
Armeniaca	*Prunus mandshurica*	WX201	Jilin, China	simple flower	41°44′49″	125°59′13″	377	SRR17479199	SRR12920660
	Prunus mume	WX206	Yunnan, China	simple flower	25°08′40″	102°44′28″	1926	SRR17479193	SRR12920640
	Prunus sibirica	WX205	Jilin, China	simple flower	41°43′42″	125°57′18″	381	SRR17479209	SRR12920641
Prunus s.str.	*Prunus ussuriensis*	WX209	Jilin, China	simple flower	41°44′49″	125°59′13″	377	SRR17479200	SRR12927898
	Prunus salicina	SN505	Shaanxi, China	simple flower	34°16′20″	108°05′04″	355	SRR17479201	SRR17543968
	Prunus cerasifera	SN506	Shaanxi, China	simple flower	34°16′20″	108°05′04″	355	SRR17479207	SRR17543970
Amygdalus	*Prunus davidiana*	WX207	Shaanxi, China	simple flower	34°07′17″	107°53′46″	975	SRR17479186	SRR12920639
	Prunus davidiana	SN501	Shaanxi, China	simple flower	34°16′20″	108°05′04″	355	SRR17479210	–
Cerasus	*Prunus maximowiczii*	WX215	Jilin, China	corymb	41°44′49″	125°59′13″	377	SRR17479204	SRR12920657
	Prunus discadenia	SN502	Shaanxi, China	raceme	33°28′60″	108°29′47″	2321	SRR17479188	SRR12927899
	Prunus serrulata	WX211	Jilin, China	corymb	41°44′49″	125°59′13″	376	SRR17479177	SRR12920636
	Prunus tomentosa	WX216	Shaanxi, China	simple flower	34°16′20″	108°05′04″	355	SRR17479198	SRR12920656
	Prunus cerasoides	WX212	Yunnan, China	corymb	25°08′40″	102°44′28″	1929	SRR17479203	SRR12927897
	Prunus japonica var. *nakaii*	SN503	Jilin, China	simple flower	41°43′42″	125°57′18″	377	SRR17479211	SRR17543972
	Prunus stipulacea	WX213	Shaanxi, China	corymb	33°23′54″	108°22′18″	2320	SRR17479178	SRR12920635
Padus	*Prunus obtusata*	WX224	Sichuan, China	raceme	30°03′15″	101°57′47″	2547	SRR17479197	SRR12920649
	Prunus virginiana	WX220	Shaanxi, China	raceme	34°16′20″	108°05′04″	355	SRR17479195	SRR12920652
	Prunus serotina	WX204	Washington DC, USA	raceme	43°54′48″	−77°02′47″	120	SRR17479181	SRR12920648
	Prunus padus	WX222	Shaanxi, China	raceme	34°07′17″	107°53′46″	1251	SRR17479174	SRR17543971
	Prunus maackii	WX221	Liaoning, China	raceme	41°46′31″	123°25′36″	26	SRR17479176	SRR12920651
	Prunus napaulensis	WX225	Sichuan, China	raceme	29°31′56″	103°20′09″	2543	SRR17479179	SRR12927896
	Prunus brachypoda	WX223	Hubei, China	raceme	31°44′40″	110°40′33″	2133	SRR17479208	SRR12920650
Lauro-cerasus	*Prunus zippeliana*	WX227	Sichuan, China	raceme	29°31′56″	103°20′09″	743	SRR17479205	SRR12920646
	Prunus laurocerasus	WX226	Rockville, Maryland, USA	raceme	38°54′48″	−77°00′47″	119	SRR17479192	SRR12920647
	Prunus jenkinsii	SN509	Yunnan, China	raceme	21°55′11″	101°16′40″	570	SRR17479194	SRR17543967
Pygeum	*Prunus topengii*	WX229	Guangdong, China	raceme	23°03′38″	113°23′41″	25	SRR17479190	SRR12920644
	Prunus arborea var. *montana*	WX228	Sichuan, China	raceme	29°31′56″	103°20′9″	745	SRR17479196	SRR12920645
	Pygeum macrocarpum	SN510	Yunnan, China	raceme	23°03′38″	113°23′41″	570	SRR17479202	SRR17543969
Maddenia	*Prunus hypoleuca*	JR324	Hubei, China	raceme	31°25′32″	110°16′51″	2134	SRR17479191	SRR13863263
	Prunus hypoleuca	JR348	Shaanxi, China	raceme	34°02′17″	107°42′12″	2813	SRR17479180	SRR13863261
	Prunus incisoserrata	JR354	Gansu, China	raceme	34°23′20″	103°55′44″	2553	SRR17479189	SRR13863259
	Prunus incisoserrata	JR440	Shaanxi, China	raceme	33°28′59″	108°29′46″	2324	SRR17479182	SRR13868097
	Prunus wilsonii	JR428	Sichuan, China	raceme	29°31′46″	103°20′07″	2940	SRR17479185	SRR13868095
	Prunus wilsonii	JR314	Shaanxi, China	raceme	33°28′60″	108°29′46″	2326	SRR17479184	SRR13868096
	Prunus hypoxantha	JR426	Sichuan, China	raceme	29°31′46″	103°20′07″	2950	SRR17479183	SRR13868094
	Prunus hypoxantha	JR374	Sichuan, China	raceme	30°03′40″	102°00′22″	2545	SRR17479187	SRR13863257
Outgroups	*Physocarpus amurensis*	WX230	Shaanxi, China	–	34°16′20″	108°05′04″	355	SRR17479206	SRR12920643
	Prinsepia uniflora	WX231	Shaanxi, China	–	34°16′20″	108°05′04″	355	SRR17479175	SRR12920642

number: NC_014697), *P. pseudocerasus* Lindl (accession number: NC_030599). and *P. yedoensis* Matsum (accession number: NC_026980). All the plastomes were successfully assembled. The seven published sequences were also used as references for annotating the plastomes with PGA (Qu et al., 2019). The annotated sequences were manually revised in Geneious v.11.0.2.

2.3. Phylogenetic analyses

Phylogenetic relationships were inferred for each of these four data sets (full sequence matrix, and those allowing a maximum of 40%, 60%, or 80% missing data per site). Concatenated Maximum Likelihood (ML) phylogenies were inferred using RAxML v.8.2.12 with 100 bootstrap replicates and 20 independent ML searches. The GTRGAMMA model of evolution was used for the full sequence matrix; the ASC_GTRGAMMA model with the Lewis ascertainment correction was used for the three SNP data sets.

Species trees were constructed using the nuclear data in a coalescent framework with SVDQuartets (Chifman and Kubatko, 2014). All quartets were evaluated for each of the four data matrices, and 100 bootstrap replicates were implemented to assess confidence in phylogenetic relationships. The 100,579 (~250 bp) loci were considered unlinked genes for the full sequence matrix; however, each SNP was considered unlinked in the three SNP data sets.

All the plastome sequences were aligned using MAFFT (Katoh and Standley, 2013) implemented in Geneious v.11.0.2 (Kearse et al., 2012). The plastome data set was used to infer the phylogeny of *Prunus* based on the maximum likelihood (ML) method. The analyses were performed by RAxML-HPC Black Box 8.2.12 (Stamatakis, 2014) with 1000 bootstrap replicates and the GTR + G model via the CIPRES Science Gateway website (Miller et al., 2010). For the plastome data set, we also reconstructed the phylogeny of *Prunus* using MrBayes v.3.2 (Ronquist et al., 2012) based on Bayesian inference (BI) methods. Bayesian inference was performed using ten million generations, the first 25% of trees were discarded as burn-in, and trees were sampled every 1000 generations.

2.4. Hybridization tests

We tested the hypothesized allopolyploid origin of the racemose group using the software package Hybrid Detector (HyDe; Blischak et al., 2018). HyDe is an approach similar to ABBA-BABA tests, and uses phylogenetic invariants arising under a coalescent model to identify hybridization among three ingroup taxa polarized by an outgroup. Within each quartet examined, the parameter γ represents the probability of one ingroup species X being sister to species Y, whereas 1-γ represents the probability of species X being sister to species Z. A significant γ value of 0.5 implies that species X is an F1 hybrid resulting from species Y and Z. We ran three separate analyses using this approach. We ran HyDe analyses exhaustively on all combinations of individual species, and then we combined accessions into three groups based on inflorescence type, i.e.,

solitary, corymbose and racemose. Additionally, we investigated hybridization within each of these three groups. *Physocarpus amurensis* (Maxim.) Maxim and *Prinsepia uniflora* Batalin were used as outgroups for all analyses.

2.5. Divergence time estimation

Divergence times were estimated using the plastome matrix of the 38 individuals in BEAST v.2.5.2 (Bouckaert et al., 2019). Based on the published fossil endocarp *Prunus wutuensis* from Shandong Province in eastern China (Li et al., 2011), a lognormal prior with an offset of 55.0 million years ago (Mya) and a standard deviation of 1 Mya were used to set the stem age of *Prunus* subg. *Cerasus*. A normal prior with offset of 35.0 Mya and a standard deviation of 2 Mya were set for the crown age of the racemose group according to the published amber of a staminate flower of *Prunus hirsutipetala* D.D. Sokoloff, Remizowa et Nuraliev from northwestern Ukraine (Sokoloff et al., 2018). A nodal calibration was set for the crown age of *Prunus* (Chin et al., 2014), using a normal prior with an offset of 61.5 Mya and a standard deviation of 3.0. The Markov chain Monte Carlo (MCMC) run was conducted for 100, 000,000 generations. Trees were sampled every 1000 generations. The log files of BEAST analysis were checked with Tracer 1.5 (Rambaut et al., 2018). Results were considered reliable once the effective sampling size (ESS) for all parameters exceeded 200. We used TreeAnnotator to generate the maximum clade credibility tree with 25% of trees discarded as burn-in.

2.6. Ancestral character state reconstruction

We traced the inflorescence evolution of *Prunus* using the BI tree inferred from the plastome data set. Three inflorescence types were defined (0) simple flowers, (1) corymbs, and (2) racemes. The detailed inflorescence information was obtained from *Flora Republicae Popularis Sinicae* (http://www.iplant.cn/frps) and observations in the field listed in Table 1. The 'Stochastic Character Mapping' method implemented in Mesquite v.3.51 (Maddison and Maddison, 2018) was employed to reconstruct ancestral character states.

3. Results

3.1. Characteristics of RAD-seq and plastome data sets

RAD-seq yielded 12.9 Gb raw data across the 38 samples included in this study. The total read number ranged from 1,436,523 to 5,977,937 among all sequenced individuals. The average Q30 percentage was 93.5% and the average GC content was 42.8% (Table S1). The SNP matrices constructed via *de novo* assembly using ipyrad consisted of a sequence matrix totaling 25,770,082 nucleotide sites per species and a SNP matrix of 1,587,718 nucleotides per species. Subsequent filtering of the assembly resulted in data matrices of 9,637, 65,176, and 448,083 sites, respectively (Table S2).

The size of *Prunus* plastomes ranged from 157,660 bp (*P. davidiana*) to 159,000 bp (*P. stipulacea* Maxim.) in length. Plastomes of all *Prunus* species we studied had a quadripartite structure, including a large single-copy (LSC, 85,764–87,692 bp) region, a small single-copy (SSC, 18,855–19,161 bp) region, and two inverted repeat (IR, 26,198–26,458 bp) regions (Table S3). The total GC content of all the *Prunus* plastomes ranged from 36.6% to 36.8%. All *Prunus* plastomes encoded 113 unique genes, including 79 protein-coding genes (CDS), four ribosomal RNAs (rRNAs), and 30 transfer RNAs (tRNAs). In addition, 17 genes were duplicated in the IRs, of which six, four, and seven were protein-coding genes, rRNAs and tRNAs, respectively (Table S3). The matrix length of the plastome sequence was 169,232 bp after trimming.

3.2. Phylogenetic relationships of Prunus

Overall, the nuclear and plastome data both showed three strongly supported major clades (Fig. 2 and Figs. S1–S6), each presenting a characteristic inflorescence architecture: the racemose group (clade 1/A), solitary-flower group (clade 2/B) and corymbose group (clade 3/C). Cytonuclear discord was not detected in the backbone of *Prunus*. Still, there were some conflicts, especially within the major clades, e.g., for section *Armeniaca* and section *Microcerasus* M. Roem. of the solitary-flower group, and within the racemose group, as well as within the corymbose group (Fig. 2). The racemose group (clade 1/A) included the subg. *Lauro-cerasus*, subg. *Padus*, and *Maddenia* and *Pygeum* groups, and its monophyly was strongly supported with a sister relationship to the clade of the remaining two major groups of *Prunus* (clades 2/B and 3/C). *Padus* species formed a monophyletic group based on ipyrad assemblies (Figs. 2, S1, S2 and S4). The nuclear phylogenies inferred using the concatenated-based method were broadly congruent with the phylogenies constructed using a coalescent-based approach (i.e., SVDQuartets) (Figs. 2 and S1–S6). However, the species trees and concatenation-based phylogenies showed a different phylogenetic placement of *Prunus serotina* Ehrh. Whereas *P. serotina* was always sister to the remaining *Padus* species in the SVDQuartets trees (Figs. 2, S1 and S2), this was the case in the concatenation-based phylogeny for only one data set—the one with a maximum of 40% missing data (Fig. S4). In the data sets with no limit on missing data (Figs. S3 and S6), *P. serotina* was sister to a clade containing all species of *Padus* and *Maddenia*. In the data set allowing up to 80% missing data, *P. serotina* was sister to *Maddenia* (Fig. S5), which matched the chloroplast topology but not the nuclear coalescent topology.

In the rest of this section, the tree generated using the ipyrad assembly allowing a maximum of 60% missing data will be used to discuss the phylogenetic relationships of *Prunus* because there were only minor differences among assemblies (Figs. 2 and S1–S6). *Lauro-cerasus* was monophyletic in none of the analyses. One species, *Prunus laurocerasus* L., was sister to all taxa from *Maddenia* and *Padus* in the nuclear tree, in contrast with the sister relationship to the *Padus* species excluding *P. serotine*, as shown in the plastome tree. Both nuclear and plastome phylogenies showed that *Maddenia* and *Pygeum* each formed a monophyletic group (Fig. 2).

The solitary-flower group (clade 2) contained not only species traditionally assigned to it (subg. *Amygdalus*, section *Armeniaca* and subg. *Prunus* s. str.), but also section *Microcerasus* species of subg. *Cerasus*. *Amygdalus* was the first diverging lineage within the solitary flower group, but there were some conflicts between the nuclear and plastome trees within clade 2. Section *Microcerasus* was a monophyletic clade in the nuclear tree but not in the plastome tree. *Armeniaca* and *Prunus* s. str. formed a sister clade, and together they were sister to *Microcerasus* in the nuclear tree. *Armeniaca* (*Prunus mandshurica* (Maxim.) Koehne and *P. mume*) formed a clade with *Microcerasus*, and one *Armeniaca* species (*Prunus sibirica* L.) was sister to *Prunus* s. str. in the plastome tree.

The corymbose group was composed of all subg. *Cerasus* species (clade 3) and was divided into two subclades, but the position of *Prunus cerasoides* Buch.-Ham. ex D. Don was not congruent in the plastome and nuclear trees. *P. cerasoides* and *P. serrulata* Lindl. formed a clade with *P. maximowiczii* that was siter to the clade of *P. stipulacea* and *P. discadenia* Koehne in the nuclear tree. *P. cerasoides* and (*P. stipulacea* + *P. discadenia*) formed a clade that was sister to the clade of *P. serrulata* and *P. maximowiczii* in the plastome tree (Fig. 2).

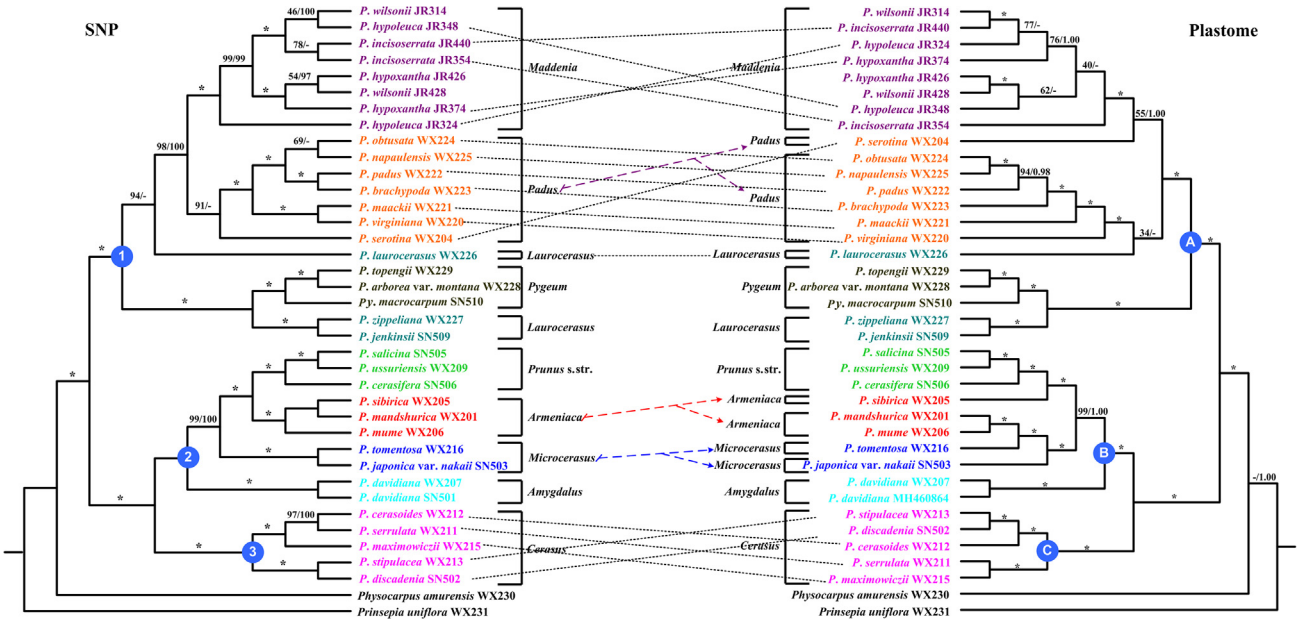


Fig. 2. The nuclear (left) and plastome (right) topologies for 36 *Prunus* samples plus two outgroups (*Physocarpus amurensis* and *Prinsepia uniflora*). SVDQuartets tree and RAxML tree (left) for 38 species constructed using ipyrad *de novo* assemblies allowing maximum 60% missing data (65,176 SNPs). Numbers at nodes indicate bootstrap support values. The support values above the branches show BS (bootstrap support) (left) and PP (posterior probability)/BS (right), and asterisks indicate 1.00/100%. Dashes represent incongruences of BI/ML tree and SVDQuartets/RAxML tree.

3.3. Hybridization tests

Of the 13,485 hypotheses of hybridization tested using HyDe, 4180 contained significant evidence of hybridization (Table S4). There was a wide range of γ values among the significant tests, and over 2/3 of γ values were less than 0.3 or greater than 0.7 (Table 2). Typically, γ values that deviate substantially from 0.5, as is the case here, indicate a more ancient history of hybridization within the group, as opposed to a first-generation hybridization event that a γ value of 0.5 would suggest. Species from the racemose group make up the majority of species included in the significant hybridization comparisons, although racemose species are often found included as either parental species or hybrids (Table 2). The HyDe analyses grouping together accessions into three groups suggested that the corymbose group may have originated via hybridization of the racemose and solitary-flower groups (Table S5). HyDe analyses of hybridization within each of the three inflorescence groups indicated high levels of hybridization (i.e., greater than 30%) (Tables S6–S8).

Table 2
The summarized significant HyDe results, representing 4180 hypotheses of hybridization tested out of a total of 13,485 investigated. The proportion of times that a species from the racemose, solitary, and corymbose group represents Parent 1 (P1), the Hybrid, or Parent 2 (P2) is summarized in the first three rows. The distribution of γ values across all significant tests is shown in the bottom row.

P1	Racemose group (66.4%)
	Solitary-flower group (22.4%)
	Corymbose group (11.2%)
Hybrid	Racemose group (47.5%)
	Solitary-flower group (32.0%)
	Corymbose group (20.5%)
P2	Racemose group (50.1%)
	Solitary-flower group (33.1%)
	Corymbose group (16.8%)
γ value	$\gamma < 0.3$ (34.0%)
	$0.3 \leq \gamma \leq 0.7$ (31.5%)
	$\gamma > 0.7$ (34.5%)

3.4. Divergence time estimation

The diversification of *Prunus* began around 67.32 Mya (95% HPD: 55.66–79.77 Mya) in the Late Cretaceous. Clades 2 and 3 split at around 58.38 Mya (95% HPD: 41.23–73.71 Mya), and diversified around 25.92 Mya (95% HPD: 12.83–58.59 Mya) and 11.45 Mya (95% HPD: 3.86–49.58 Mya), respectively (Fig. 3). Clade 1 diversified around 35.41 Mya (95% HPD: 26.51–43.39 Mya). The divergence of *Pygeum* and *Maddenia* groups in clade 1 occurred at 12.24 Mya (95% HPD: 2.72–29.29 Mya) and 2.81 Mya (95% HPD: 0.86–12.82 Mya), respectively.

3.5. Ancestral character state reconstruction

Our results indicated that the ancestral inflorescence type of *Prunus* was the raceme. The ancestral state of clades 2 and 3 in *Prunus* was a simple flower. As such, the simple flower and corymb states were derived. In the corymbose group, there was an evolutionary reversal to raceme in *P. discadenia* SN502 (Fig. 4).

4. Discussion

Below, we discuss the main findings of our phylogenetic results and inferences of hybridization events, divergence times, and inflorescence evolution in *Prunus*.

4.1. Phylogenetic relationships

Our study integrated nuclear and plastome data to infer phylogenetic relationships of *Prunus* based on species sampling representing all major lineages within the genus. Data sets from both nuclear and plastid genomes supported that *Prunus* consisted of three main clades—the racemose group, the solitary-flower

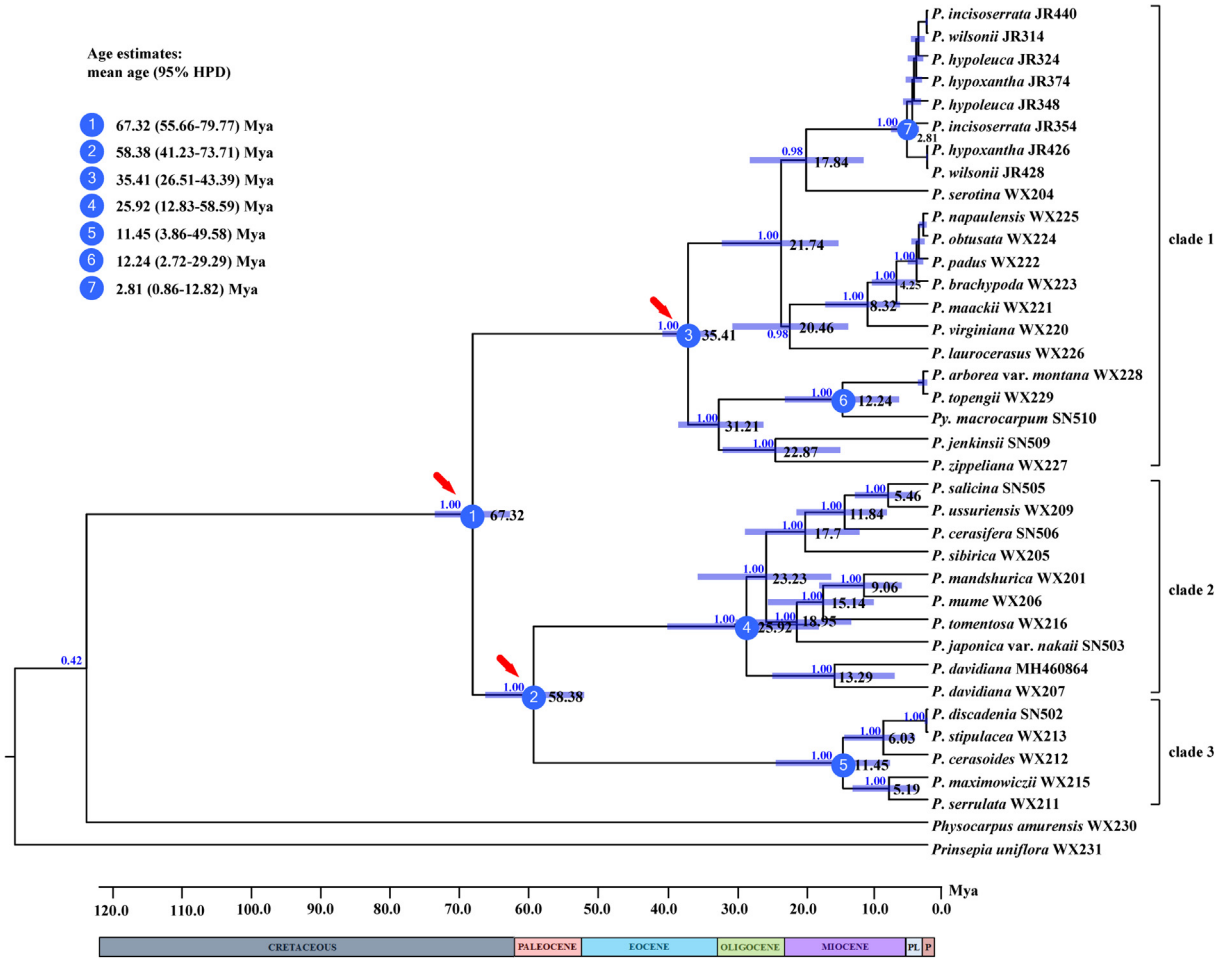


Fig. 3. Chronogram of *Prunus* based on plastome data sets inferred from BEAST. Blue bars represent the 95% high posterior density credibility interval for node ages. Three calibration points are indicated with red arrows. Nodes of interests were marked as 1–7. Bayesian posterior probabilities are given above branches.

group, and the corymbose group. The corymbose group was sister to the solitary-flower group, and then together sister to the racemose group with strong support. This topology was consistent with previous findings based on plastid genes (Chin et al., 2010, 2014; Zhao et al., 2016) and those from some single-copy nuclear genes and plastomes (Hodel et al., 2021), but in conflict with results from other analyses of nuclear genes (Chin et al., 2010, 2014). Although our two data sets resulted in a consistently robust backbone, strong phylogenetic conflict was recovered at some shallow-level nodes within each of the three main groups.

The racemose group was previously considered paraphyletic according to phylogenetic trees using a few nuclear gene loci (Chin et al., 2014; Zhao et al., 2016), but it was monophyletic in both nuclear and plastome trees (Fig. 2 and Hodel et al., 2021). The observed monophyly of the racemose group may have resulted from the majority of nuclear loci included in this study tracking the maternal phylogeny of the racemose group. Often, a few outlier genes with outsized influence may skew phylogenetic relationships, so even when relationships appear strongly supported, underlying cytonuclear or gene tree discord may give false confidence in incorrect relationships (Walker et al., 2022). The uncertain phylogenetic placement of the racemose species *P. serotina*, with one concatenation phylogeny matching the plastid topology as opposed to the nuclear coalescent topology, may indicate that the maternal phylogeny readily dominated some of the SNP data sets when constructing concatenated trees (Fig. S5). In any case, the origin of the racemose group needs to be better explored with single-copy nuclear data that track multiple homologous loci of a specific gene (Zimmer and Wen, 2015) and/or whole nuclear genome data. Both *Pygeum* and *Maddenia* groups were strongly supported as monophyletic, which is consistent with previous results (Zhao et al., 2018; Su et al., 2021). The paraphyly of subg. *Padus* and subg. *Lauro-cerasus* was suggested in previous studies (Bortiri et al., 2002; Wen et al., 2008; Chin et al., 2010, 2014; Liu et al., 2013; Shi et al., 2013; Zhao et al., 2016, 2018). Our results based on plastome data further confirm these conclusions. *Padus* and *Lauro-cerasus* may have each derived from multiple hybridization and allopolyploidy events (Zhao et al., 2016). However, the three trees based on ipyrad *de novo* assemblies show a monophyletic *Padus*, and the difference in the position of *P. serotina* was the most obvious incongruence among the topologies obtained with *de novo* assembly and those published in previous studies (Liu et al., 2013; Chin et al., 2014; Zhao et al., 2018) in nuclear trees.

Prunus maackii Rupr., a species mainly distributed in northeast China, Korea, and eastern Russia, has been treated as a member of subg. *Padus* (Rehder, 1956; Yü et al., 1986; Lu et al., 2003). This species has racemose and deciduous inflorescences similar to some species of *Padus* (Rehder, 1956; Wen et al., 2008). Yet it resembles some members of *Cerasus* with highly incised stipules (Liu et al., 2013) and a short inflorescence. *P. maackii* has been supported to be nested within subg. *Cerasus* based on nuclear and plastid sequences (Lee and Wen, 2001; Bortiri et al., 2006; Wen et al., 2008; Shi et al., 2013; Liu et al.,

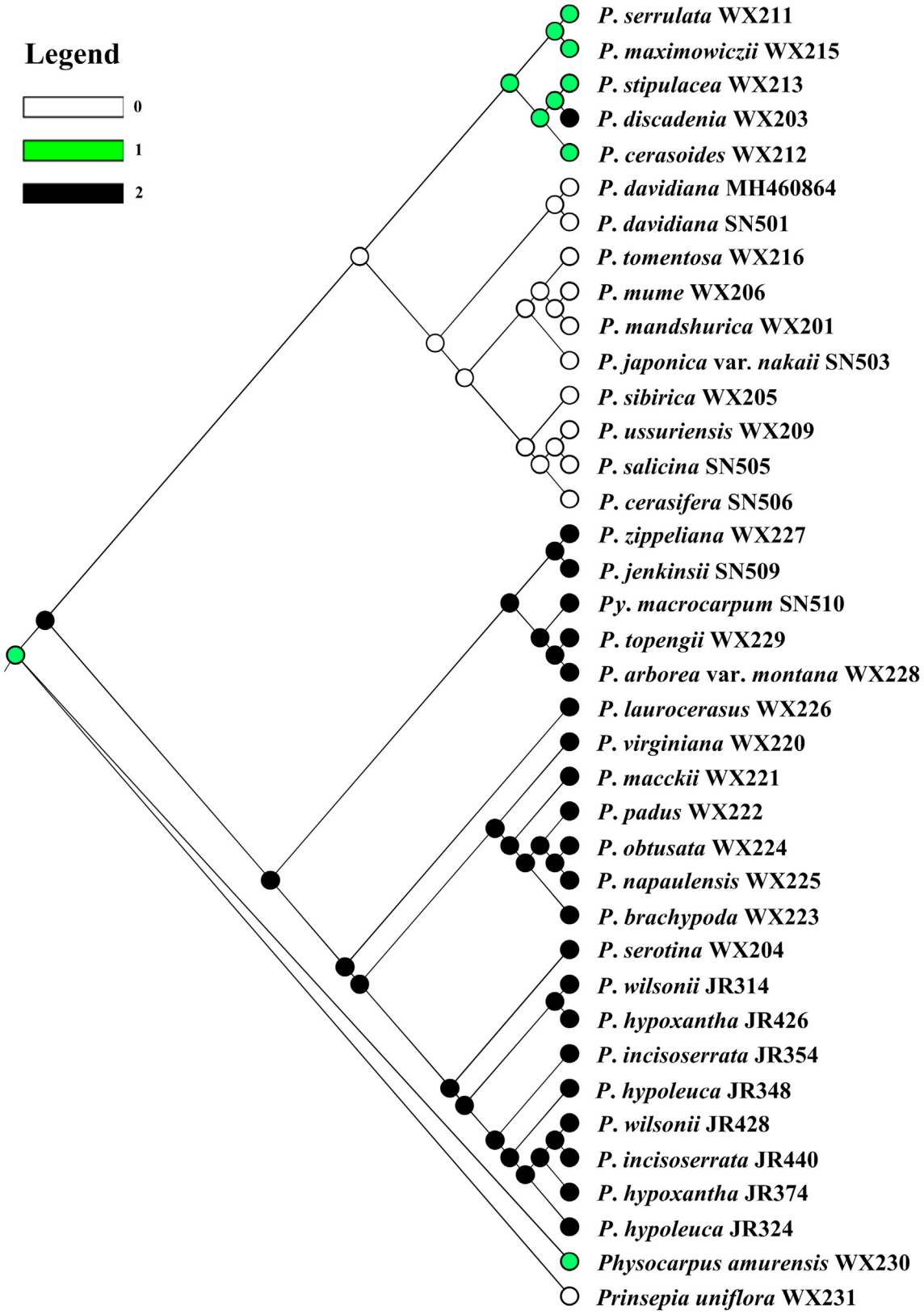


Fig. 4. Ancestral character reconstruction of *Prunus* inflorescences using the Stochastic Character Mapping implemented in the program Mesquite.

2013; Chin et al., 2014; Zhao et al., 2016). This species was reported to hybridize with *P. maximowiczii* of the *Cerasus* group and formed natural hybrids (Wen et al., 2008; Liu et al., 2013; Shi et al., 2013). Thus, some researchers suggested *P. maackii* be placed in subg. *Cerasus* (Wen et al., 2008; Liu et al., 2013). Yet, we found that it formed a clade with *Padus* species in both nuclear and plastome trees with high support for the first time (Fig. 2), which challenges previous studies. Nevertheless, *P. maackii* is similar to most members of subg. *Padus*, and they share two morphological characteristics, i.e., terminal racemes with a few leaves at the base of their peduncle and leaves without glands on the petiole. Our results hence support the placement of *P. maackii* in subg. *Padus*.

The solitary-flower group showed several conflicts between plastome and nuclear trees (Fig. 2). Section *Armeniaca* was monophyletic in the nuclear tree, but not in the plastome tree. Due to maternal inheritance of plastomes and biparental inheritance of nuclear data, section *Armeniaca* may be derived from one paternal parent and several maternal lineages.

Even though the corymbose group refers to all species of subg. *Cerasus*, species of *Microcerasus* were scattered in the solitary-flower group in our analyses. *Cerasus*, excluding section *Microcerasus*, was monophyletic and had some cytonuclear discord among *Cerasus* species (Fig. 2). *Microcerasus* was regarded as a section of subgenus *Cerasus* by Rehder (1956) and as a subgenus of the more narrowly defined genus *Cerasus* by Yü et al. (1986). *Microcerasus* comprises shrubby and woody species with three axillary buds and a short pedicel, similar to the solitary-flower group members (Lee and Wen, 2001; Wen et al., 2008; Shi et al., 2013). However, true cherries share only one axillary bud at each leaf axil. Moreover, it was reported that *Microcerasus* species more readily hybridized with species of the solitary-flower group than with those of the *Cerasus* group (Kataoka et al., 1988).

Two decades ago, Lee and Wen (2001) were the first to systematically investigate the phylogenetic relationships of *Prunus* based on ITS sequences, and showed that two *Microcerasus* species had been closely related to the solitary-flower group rather than to the corymbose group. This result was supported by those inferred from different molecular markers in many subsequent studies (Bortiri et al., 2001, 2002, 2006; Wen et al., 2008; Liu et al., 2013; Shi et al., 2013; Chin et al., 2014; Zhao et al., 2016). Our nuclear and plastid results revealed that the *Microcerasus* species (*Prunus tomentosa* Thunb. and *P. japonica* var. *nakaii* (H. Lév.) Rehder) were nested within the solitary-flower group, which is congruent with previous studies. Our analysis did not support the inclusion of *Microcerasus* into *Cerasus*. We found that the ovule developmental characteristics of *P. tomentosa* were more similar to members of the solitary-flower group than to *Cerasus* (personal observation by L. Zhao). Shi et al. (2013) also suggested that the *Cerasus* group only included true cherries, and that *Microcerasus* did not belong to it. Based on morphological and molecular evidence, it seemed reasonable to assign *Microcerasus* to the solitary-flower group. Mowrey and Werner (1990) pointed out that due to its paraphyly in all analyses *Microcerasus* was not a clade, which has been supported by subsequent studies (Lee and Wen, 2001; Bortiri et al., 2002; Wen et al., 2008; Chin et al., 2014). Our results also show that *Microcerasus* species did not form a clade in the plastome tree. Therefore, further studies should sample additional *Microcerasus* species and closely related taxa to better understand their origin(s) and systematic placement.

4.2. *On the hybrid origins of major* Prunus *lineages*

Zhao et al. (2016) hypothesized that the racemose group derived from multiple hybridization events and allopolyploid speciation events. The maternal parent may have been extinct, and the paternal parents shared common ancestry with various members of the corymbose-solitary flower lineages, leading to the observed conflicts between the plastid and nuclear topologies (Zhao et al., 2016). However, our results support the monophyly of the racemose group (Fig. 2).

Although these results challenge the hypothesis of multiple hybrid origins of the racemose group, the HyDe analysis based on RAD-seq data detected frequent hybridization events in this group. In concert with previous studies, the HyDe results indicate pervasive hybridization and/or allopolyploidy within *Prunus*, especially within the racemose group (Table 2). Our goal was to test the hypothesized allopolyploid origin of the racemose group, but it is challenging to use hybrid detection methods based on extant taxa to assess ancient hybridization that may have involved extinct lineages. Furthermore, subsequent introgression and/or repeated hybridization can lead to conflicting genomic signals in different species. Nonetheless, a large number of hybridization events detected by HyDe highlights the frequency of hybridization within this genus. Analyses of individual species and within groups indicate that hybridization was pervasive throughout *Prunus*, and our data do not readily confirm any specific hypotheses of ancient hybridization (Tables S5–S8). Because many species in the racemose group are polyploid, HyDe may pick up signals of genome doubling (i.e., allopolyploidy) instead of homoploid hybridization.

Previous studies with less representation of racemose species (Hodel et al., 2021) found that approximately 1% of hybridization tests were significant, in contrast with the present study in which 31% of the tests were significant. A similar approach in another Rosaceae lineage with frequent hybridization (the apple genus *Malus*) found that 24% of hybridization tests were significant (Liu et al., 2022a). A study of the Hawaiian genus *Myrsine* using similar markers as the present study (i.e., RAD-seq) found similarly high levels of hybridization using HyDe (between 27.5% and 30.4% of tests significant, depending on the data set; Appelhans et al., 2020). RAD-seq data have been used successfully for non-polyploid taxa with fewer reticulation events (e.g., Hodel et al., 2022). However, RAD-seq loci cannot be used for accurately estimating ancient hybridization events between the polyploid species and its close diploid relatives (e.g., Wang et al., 2021), especially in the presence of extinction events, genome doubling and/or subsequent recent hybridizations.

Our HyDe analyses with accessions grouped based on inflorescence type showed a potential signal of ancient hybridization of the corymbose group, i.e., between the racemose group and solitary-flower group (Table S5), in contrast to the hypothesis of allopolyploid origins of the racemose group suggested by previous studies (Chin et al., 2014; Zhao et al., 2016). While homoploid hybrid origin of the corymbose group with a maternal 'Solitary flower' lineage and a diploid paternal ancestral 'Racemose' lineage is plausible (Chin et al., 2014), it is not supported by the higher ploidy levels of all extant members of the racemose group, nor is it consistent with patterns of cytonuclear discord concerning the relationships among the three major groups observed in previous studies (Chin et al., 2014; Zhao et al., 2016) or the lack of cytonuclear discord concerning those relationships observed here. Moreover, our HyDe analyses using extant taxa grouped based on inflorescence types could not explicitly test Zhao et al.'s (2016) hypothesis of multiple allopolyploid origins of the racemose group, because that hypothesis invoked members of a now-extinct lineage as the maternal parents in the hybridization events. In fact, if an ancestral member of the corymbose group acted as one or more of paternal parents in the hybridization events, as suggested by Zhao et al. (2016), then one would expect exactly the results seen in our species group HyDe tests here, because some nuclear genes would group the corymbose species with their sister solitary-flower group and others would group them with the hybrid racemose lineages to

which they gave rise via hybridization. On the other hand, the corymbose inflorescence seems to represent an intermediate morphological state between the raceme and the simple flowers, which may be consistent with the above hybridization hypothesis of the corymbose group. Future additional studies are needed to rigorously test these competing hypotheses using phylogenomic and developmental morphological data.

4.3. *Temporal diversification of main clades in* Prunus

A previous study based on four plastid genes and ITS sequences estimated that the ancestor of *Prunus* emerged between 56.7 and 67.4 Mya (Chin et al., 2014). The age of *Prunus* in the present study slightly predates their estimate (Fig. 3). In our chronogram, the ancestor of clades 2 and 3 firstly split around 58.38 Mya (95% HPD: 41.23–73.71 Mya). This is congruent with the fossil record of *P. wutuensis*, which mostly resembled extant *P. yedoensis* of subgenus *Cerasus* (Li et al., 2011). The evolution of plants has often been impacted by paleoclimatic and geologic events. During the boundaries between epochs of the Paleogene, key paleoclimatic events caused multiple effects on the evolution and distribution of plants. Around 55.8 Mya, there was an abrupt period of global warming caused by a transient burst of CO_2, known as the Paleocene-Eocene Thermal Maximum (PETM) (Currano et al., 2008; McInerney and Wing, 2011). During this period, floristic changes have been shown to occur in response to climate fluctuations (Wing and Currano, 2013). Clades 2 and 3 split around 58.38 Mya, which corresponds to the PETM. Therefore, our results seem to indicate that climatic change may have influenced the diversification of *Prunus* during the PETM.

In addition, the Eocene–Oligocene transition (EOT, 30–40 Mya) in the Cenozoic witnessed a global cooling and led to a reorganization of organisms (Prothero, 1994; Sun et al., 2014; Deng et al., 2020). The diversification of clades 1 and 2 was herein estimated to be around 35.41 Mya (95% HPD: 26.51–43.39 Mya) in the Late Eocene and 25.92 Mya (95% HPD: 12.83–58.59 Mya) in the Late Oligocene, respectively (Fig. 3). There was another climatic cooling event after the Middle Miocene Climatic Optimum (MMCO, around 15 Mya) (Flower and Kennett, 1994). The molecular dating analysis indicate that clade 3 of *Prunus* initially diverged when a climatic cooling event occurred after MMCO. The divergence time of the three clades coincided with the two cooling events, which also indicates that the paleoclimatic events might have impacted *Prunus* diversification and evolution.

4.4. *Ancestral character state reconstruction of inflorescences*

By tracking the evolution of inflorescence types on the *Prunus* phylogeny, we inferred that the raceme was the ancestral state of inflorescence in this genus. The solitary flower and corymbose inflorescence types might be derived from the reduction of the flower number and suppression of the rachis, respectively. This result is consistent with the hypothesis that racemes are a primitive state in angiosperms (Stebbins, 1973). Although terminal solitary flowers (Parkin, 1914) and panicles (Takhtajan, 1991) have also been regarded as an ancestral state of inflorescences, there has been some skepticism about the reliability of any of these ideas because the earliest-diverging angiosperms, the ANITA grade, do not possess any panicles, and single flowers terminal on shoots are only present in several genera, such as *Austrobaileya* C.T. White (in part), *Schisandra* Michx (in part), and *Illicium* L. (Endress, 2010).

Simple flowers were the ancestral state of clades 2 and 3 of *Prunus* based on mature inflorescence types. In fact, a lateral floral primordium was initiated but it did not develop further, contributing to simple flowers at maturity in *P. persica* (personal observation by L. Zhao). *P. salicina* and *P. americana* Marshall have a typical inflorescence type with 2–3 flowers clustered in a bud (personal observation by L. Zhao and J. Wen), which resemble a corymb more than a simple flower. Such developmental/morphological evidence suggests that there may be a transition between simple flowers and corymbs. Hence, we assume the ancestral state of clades 2 and 3 had a transitional inflorescence type, which shows simple flowers only at maturity, with more than one flower observed even then. Corymb and simple flowers in *Prunus* were formed by continuing and stopping development of lateral floral primordium from their ancestor, respectively. Furthermore, *P. discadenia* SN502, a member of subg. *Cerasus*, had racemose inflorescence. It seems that this represents a case of reverse evolution in *Prunus*, i.e., racemes likely evolved from corymbs by elongation of the rachis. However, we only analyzed the mature character state of *Prunus* and did not provide robust evidence for developmental evolutionary histories of inflorescences in this study. Future studies integrating inflorescence development and phylogenetic analyses are needed to understand the evolutionary directionality of *Prunus* inflorescences better.

5. Conclusions

This study sheds light on the phylogenetic relationships, hybridization events and inflorescence evolution of *Prunus*, an economically important plant group. RAD-seq data have been successfully used to elucidate the phylogenetic relationships of various polyploid lineages, such as bamboos (Guo et al., 2021) and *Salix* L. (Wagner et al., 2020). However, RAD-seq data cannot distinguish orthologs from paralogs for the polyploid species, limiting their utility for tracing ancient hybridization events. Thus, additional studies using whole-genome sequencing or genome resequencing including more extensive sampling of the racemose group and their allies are needed to test the competing hypotheses of multiple allopolyploid origins of the racemose group vs. the homoploid hybrid origin of the corymbose group (Zhao et al., 2016).

Authors contributions

Na Su: Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. **Richard G. J. Hodel**: Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. **Xi Wang**: Formal analysis, Investigation, Methodology, Software, Visualization, Writing- original draft. **Jun-ru Wang**: Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft. **Si-yu Xie**: Methodology, Software, Visualization, Writing - original draft. **Chao-xia Gui**: Writing - original draft. **Ling Zhang**: Writing - original draft. **Zhao-yang Chang**: Writing - original draft. **Liang Zhao**: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing - review & editing. **Daniel Potter**: Conceptualization, Investigation, Methodology, Project administration, Supervision, Validation, Writing - review & editing. **Jun Wen**: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, 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 thank Prof. Zhong-Hu Li and Dr. Miao–Miao Ju of Northwest University and Dr. Li Feng of Xi'an Jiaotong University, Dr. Fu-Zhen Guo and Xiao-Hua He from Instrument Sharing Platform of Northwest A&F University and Molecular Biology Experiment Center, Germplasm Bank of Wild Species in Southwest China for help with data analysis, and Mr Hai-Ning Li, Xiao-Jie Li and Prof. You Zhou for providing plant photographs (Fig. 1L–O). We also thank Dr. Florian Jabbour and Dr. Valéry Malécot for their valuable comments. Computational analyses were partially conducted on the Smithsonian Institution High Performance Computing Cluster (SI/HPC, "Hydra": https://doi.org/10.25572/SIHPC). This work was supported by grants from National Natural Science Foundation of China (No. 32170381 and 31770200).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pld.2023.03.013.

References

Andro, M.C., Riffaud, J.P., 1995. *Pygeum africana* extract for the treatment of patients with benign prostatic hyperplasia: a review of 25 years of published experience. Curr. Ther. Res. 56, 796–817.
Appelhans, M.S., Paetzold, C., Wood, K.R., et al., 2020. RADseq resolves the phylogeny of Hawaiian *Myrsine* (Primulaceae) and provides evidence for hybridization. J. Syst. Evol. 58, 823–840.
Bai, H.R., Oyebanji, O., Zhang, R., et al., 2021. Plastid phylogenomic insights into the evolution of subfamily Dialioideae (Leguminosae). Plant Divers. 41, 27–34.
Benlloch, R., Berbel, A., Serrano-Mislata, A., et al., 2007. Floral initiation and inflorescence architecture: a comparative view. Ann. Bot. 100, 659–676.
Blischak, P.D., Chifman, J., Wolfe, A.D., et al., 2018. HyDe: a Python package for genome-scale hybridization detection. Syst. Biol. 67, 821–829.
Bortiri, E., Oh, S.H., Jiang, J.G., et al., 2001. Phylogeny and systematics of *Prunus* (Rosaceae) as determined by sequences analysis of ITS and the chloroplast *trnL-trnF* spacer DNA. Syst. Bot. 26, 797–807.
Bortiri, E., Oh, S.H., Gao, F.Y., et al., 2002. The phylogenetic utility of nucleotide sequences of sorbitol 6-phosphate dehydrogenase in *Prunus* (Rosaceae). Am. J. Bot. 89, 1697–1708.
Bortiri, E.S., Vanden, B., Potter, D., 2006. Phylogenetic analysis of morphology in *Prunus* reveals extensive homoplasy. Plant Syst. Evol. 259, 53–71.
Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., et al., 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15, e1006650.
Chifman, J., Kubatko, L., 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics 30, 3317–3324.
Chin, S.W., Wen, J., Johnson, G., et al., 2010. Merging *Maddenia* with the morphologically diverse *Prunus* (Rosaceae). Bot. J. Linn. Soc. 163, 236–245.
Chin, S.W., Shaw, J., Haberle, R., et al., 2014. Diversification of almonds, peaches, plums and cherries–molecular systematics and biogeographic history of *Prunus* (Rosaceae). Mol. Phylogenet. Evol. 76, 34–38.
Currano, E.D., Wilf, P., Wing, S.L., et al., 2008. Sharply increased insect herbivory during the paleocene–eocene thermal maximum. Proc. Natl. Acad. Sci. U.S.A. 105, 1960–1964.
Danacek, P., Auton, A., Abecasis, G., et al., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158.
Deng, W.Y.D., Su, T., Wappler, T., et al., 2020. Sharp changes in plant diversity and plant-herbivore interactions during the Eocene–Oligocene transition on the southeastern Qinghai-Tibetan Plateau. Global Planet. Change 194, 103293.
Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.
Eaton, D.A.R., 2014. PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. Bioinformatics 30, 1844–1849.
Eaton, D.A.R., Overcast, I., 2020. Ipyrad: interactive assembly and analysis of RADseq datasets. Bioinformatics 36, 2592–2594.
Endress, P.K., 2010. Disentangling confusions in inflorescence morphology: patterns and diversity of reproductive shoot ramification in angiosperms. J. Syst. Evol. 48, 225–239.
Flower, B.P., Kennett, J.P., 1994. The middle Miocene climatic transition: east Antarctic ice sheet development, deep ocean circulation and global carbon cycling. Palaeogeogr. Palaeoclimatol. Palaeoecol. 108, 537–555.
Gerrath, J.M., Posluszny, U., Ickert-Bond, S.M., et al., 2017. Inflorescence morphology and development in the basal rosid lineage Vitales. J. Syst. Evol. 55, 542–558.
Guo, C., Ma, P.F., Yang, G.Q., et al., 2021. Parallel ddRAD and genome skimming analyses reveal a radiative and reticulate evolutionary history of the temperate bamboos. Syst. Biol. 70, 756–773.
Hodel, R.G.J., Zimmer, E., Wen, J., 2021. A phylogenomic approach resolves the backbone of *Prunus* (Rosaceae) and identifies signals of hybridization and allopolyploidy. Mol. Phylogenet. Evol. 160, 107118.
Hodel, R.G.J., Massatti, R., Knowles, L., 2022. Hybrid enrichment of adaptive variation revealed by genotype-environment associations in montane sedges. Mol. Ecol. 31, 3722–3737.
Jin, J.J., Yu, W.B., Yang, J.B., et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. Genome Biol. 21, 241.
Kalkman, C., 2004. Rosaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants VI. Springer, Berlin, pp. 343–386.
Kataoka, I., Sugiura, A., Tomana, T., 1988. Interspecific hybridization between *Microcerasus* and other *Prunus spp*. J. Jpn. Soc. Hortic. Sci. 56, 398–407.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvement in performance and usability. Mol. Biol. Evol. 30, 772–780.
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.
Lee, S., Wen, J., 2001. A phylogenetic analysis of *Prunus* and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. Am. J. Bot. 88, 150–160.
Lee, S.Y., Xu, K.W., Huang, C.Y., et al., 2022. Molecular phylogenetic analyses based on the complete plastid genomes and nuclear sequences reveal *Daphne* (Thymelaeaceae) to be non-monophyletic as current circumscription. Plant Divers. 44, 279–289.
Lei, F.W., Tong, L., Zhu, Y.X., et al., 2021. Plastid phylogenomics and biogeography of the medicinal plant lineage Hyoscyameae (Solanaceae). Plant Divers. 43, 192–197.
Li, Y., Smith, T., Liu, C.J., et al., 2011. Endocarps of *Prunus* (Rosaceae: prunoideae) from the early Eocene of wutu, Shandong Province, China. Taxon 60, 555–564.
Li, H.T., Yi, T.S., Gao, L.M., et al., 2019. Origin of angiosperms and the puzzle of the Jurassic gap. Nat. Plants 5, 461–470.
Li, Q.J., Su, N., Zhang, L., et al., 2020. Chloroplast genomes elucidate diversity, phylogeny, and taxonomy of *Pulsatilla* (Ranunculaceae). Sci. Rep. 10, 19781.
Li, H.T., Luo, Y., Gan, L., et al., 2021a. Plastid phylogenomic insights into relationships of all flowering plant families. BMC Biology 19, 232.
Li, X.P., Zhao, Y.M., Tu, X.D., et al., 2021b. Comparative analysis of plastomes in Oxalidaceae: phylogenetic relationships and potential molecular markers. Plant Divers. 43, 281–291.
Liu, X.L., Wen, J., Nie, Z.L., et al., 2013. Polyphyly of the *Padus* group of *Prunus* (Rosaceae) and the evolution of biogeographic disjunctions between eastern Asia and eastern North America. J. Plant Res. 126, 351–361.
Liu, B.B., Hong, D.Y., Zhou, S.L., et al., 2019. Phylogenomic analyses of the *Photinia* complex support the recognition of a new genus *Phippsiomeles* and the resurrection of a redefined *Stranvaesia* in Maleae (Rosaceae). J. Syst. Evol. 57, 678–694.
Liu, B.B., Campbell, C.S., Hong, D.Y., et al., 2020a. Phylogenetic relationships and chloroplast capture in the *Amelanchier-Malacomeles-Peraphyllum* clade (Maleae, Rosaceae): evidence from chloroplast genome and nuclear ribosomal DNA data using genome skimming. Mol. Phylogenet. Evol. 147, 106784.
Liu, B.B., Liu, G.N., Hong, D.Y., et al., 2020b. *Eriobotrya* belongs to *Rhaphiolepis* (Maleae, Rosaceae): evidence from chloroplast genome and nuclear ribosomal DNA data. Front. Plant Sci. 10, 1731.
Liu, B.B., Ma, Z.Y., Ren, C., et al., 2021. Capturing single-copy nuclear genes, organellar genomes, and nuclear ribosomal DNA from deep genome skimming data for plant phylogenetics: a case study in Vitaceae. J. Syst. Evol. 59, 1124–1138.
Liu, B.B., Ren, C., Kwak, M., et al., 2022a. Phylogenomic conflict analyses in the apple genus *Malus* s.l. reveal widespread hybridization and allopolyploidy driving diversification, with insights into the complex biogeographic history in the Northern Hemisphere. J. Integr. Plant Biol. 64, 1020–1043.
Liu, C., Chen, H.H., Tang, L.Z., et al., 2022b. Plastid genome evolution of a monophyletic group in the subtribe Lauriineae (Laureae, Lauraceae). Plant Divers. 44, 377–388.
Lu, L.D., Gu, C.Z., Li, C.L., et al., 2003. Rosaceae. In: Wu, Z.Y., Raven, P.H., Hong, D.Y. (Eds.), Flora of China. Beijing, Science Press. Missouri Botanical Garden Press, St. Louis, MO, pp. 46–434.
Ma, Q., Zhang, W.H., Xiang, Q.Y., 2017. Evolution and developmental genetics of floral display-a review of progress. J. Syst. Evol. 55, 487–515.
Ma, Z.Y., Wen, J., Tian, J.P., et al., 2018. Testing reticulate evolution of four *Vitis* species from East Asia using restriction-site associated DNA sequencing. J. Syst. Evol. 56, 311–399.
Maddison, W.P., Maddison, D.R., 2018. Mesquite: a Modular System for Evolutionary Analysis v3.51 (Version 3.51). http://www.mesquiteproject.org.
McInerney, F.A., Wing, S.L., 2011. The Paleocene–Eocene Thermal Maximum: a perturbation of the carbon cycle, climate, and biosphere with implications for the future. Annu. Rev. Earth Planet Sci. 39, 489–516.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE). New Orleans, pp. 1–8.
Mowrey, B.D., Werner, D.J., 1990. Phylogenetic relationships among species of Prunus as inferred by isozyme markers. Theor. Appl. Genet. 80, 129–133.
Mu, X.Y., Tong, L., Sun, M., et al., 2020. Phylogeny and divergence time estimation of the walnut family (Juglandaceae) based on nuclear RAD-Seq and chloroplast genome data. Mol. Phylogenet. Evol. 147, 106802.
Parkin, J., 1914. The evolution of the inflorescence. Bot. J. Linn. Soc. 42, 511–563.

Prothero, D.R., 1994. The late Eocene-Oligocene extinctions. Annu. Rev. Earth Planet Sci. 22, 145–165.
Qu, X.J., Moore, M.J., Li, D.Z., et al., 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods 15, 1–12.
Rambaut, A., Drummond, A.J., Xie, D., et al., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904.
Rehder, A., 1956. Manual of Cultivated Trees and Shrubs Hardy in North America Exclusive of the Subtropical and Warmer Temperate Regions, second ed. Macmillan, New York.
Ronquist, F., Teslenko, M., van der Mark, P., et al., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Shi, S., Li, J.L., Sun, J.H., et al., 2013. Phylogeny and classification of *Prunus* sensu lato (Rosaceae). J. Integr. Plant Biol. 55, 1069–1079.
Sokoloff, D.D., Ignatov, M.S., Remizowa, M.V., et al., 2018. Staminate flower of *Prunus s.l.* (Rosaceae) from Eocene rovno amber (Ukraine). J. Plant Res. 131, 925–943.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stebbins, G.L., 1973. Evolutionary trends in the inflorescence of angiosperms. Flora 162, 501–528.
Su, N., Liu, B.B., Wang, J.R., et al., 2021. On the species delimitation of the *Maddenia* group of *Prunus* (Rosaceae): evidence from plastome and nuclear sequences and morphology. Front. Plant Sci. 12, 743643.
Sun, X.W., Liu, D.Y., Zhang, X.F., et al., 2013. SLAF-seq: an efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. PLoS One 8, e58700.
Sun, J.M., Ni, X.J., Bi, S.D., et al., 2014. Synchronous turnover of flora, fauna, and climate at the Eocene–Oligocene boundary in Asia. Sci. Rep. 4, 7463.
Takhtajan, A., 1991. Evolutionary Trends in Flowering Plants. Columbia University Press, New York.
Thode, V.A., Lohmann, L.G., Sanmartín, I., 2020. Evaluating character partitioning and molecular models in plastid phylogenomics at low taxonomic levels: a case study using *Amphilophium* (Bignonieae, Bignoniaceae). J. Syst. Evol. 58, 1071–1089.
Vargas, O.M., Ortiz, E.M., Simpson, B.B., 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). New Phytol. 214, 1736–1750.
Wagner, N.D., He, L., Hörandl, E., 2020. Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. Front. Plant Sci. 11, 1077.
Walker, J.F., Smith, S.A., Hodel, R.G.J., et al., 2022. Concordance-based approaches for the inference of relationships and molecular rates with phylogenomic data sets. Syst. Biol. 71, 943–958.
Wang, Y.B., Liu, B.B., Nie, Z.L., et al., 2020. Major clades and a revised classification of *Magnolia* and Magnoliaceae based on whole plastid genome sequences via genome skimming. J. Syst. Evol. 58, 673–695.
Wang, N., Kelly, L.J., McAllister, H.A., et al., 2021. Resolving phylogeny and polyploid parentage using genus-wide genome-wide sequence data from birch trees. Mol. Phylogenet. Evol. 160, 107126.
Wen, J., Berggren, S.T., Lee, C.H., et al., 2008. Phylogenetic inferences in *Prunus* (Rosaceae) using chloroplast *ndhF* and nuclear ribosomal ITS sequences. J. Syst. Evol. 46, 322–332.
Wing, S.T., Currano, E.D., 2013. Plant response to a global greenhouse event 56 million years ago. Am. J. Bot. 100, 1234–1254.
Wu, H., Ma, P.F., Li, H.T., et al., 2021. Comparative plastomic analysis and insights into the phylogeny of *Salvia* (Lamiaceae). Plant Divers. 43, 15–26.
Xiang, Y.Z., Huang, C.H., Hu, Y., et al., 2016. Well-resolved Rosaceae nuclear phylogeny facilitates geological time and genome duplication analyses and ancestral fruit character reconstruction. Mol. Biol. Evol. 34, 262–281.
Xu, Y.L., Shen, H.H., Du, X.Y., et al., 2022. Plastome characteristics and species identification of Chinese medicinal wintergreens (*Gaultheria*, Ericaceae). Plant Divers. 44, 519–529.
Yü, T.T., Lu, L.T., Ku, T.C., et al., 1986. Rosaceae (3) Prunoideae. In: Yü, T.T. (Ed.), Flora Reipublicae Popularis Sinicae, vol. 38. Science Press, Beijing, pp. 1–133.
Zhang, N., Wen, J., Zimmer, E.A., 2015. Congruent deep relationships in the grape family (Vitaceae) based on sequences of chloroplast genomes and mitochondrial genes via genome skimming. PLoS One 10, e0144701.
Zhang, S.D., Jin, J.J., Chen, S.Y., et al., 2017. Diversification of Rosaceae since the late cretaceous based on plastid phylogenomics. New Phytol. 214, 1355–1367.
Zhang, Y.M., Han, L.J., Yang, C.W., et al., 2022. Comparative chloroplast genome analysis of medicinally important *Veratrum* (Melanthiaceae) in China: insights into genomic characterization and phylogenetic relationships. Plant Divers. 44, 70–82.
Zhao, L., Jiang, X.W., Zuo, Y.J., et al., 2016. Multiple events of allopolyploidy in the evolution of the racemose lineages in *Prunus* (Rosaceae) based on integrated evidence from nuclear and plastid data. PLoS One 11, e0157123.
Zhao, L., Potter, D., Xu, Y., et al., 2018. Phylogeny and spatio-temporal diversification of *Prunus* subgenus *Laurocerasus* section *Mesopygeum* (Rosaceae) in the Malesian region. J. Syst. Evol. 56, 637–651.
Zhou, M.M., Yang, G.Q., Sun, G.L., et al., 2020. Resolving complicated relationships of the *Panax bipinnatifidus* complex in southwestern China by RAD-seq data. Mol. Phylogenet. Evol. 149, 106851.
Zimmer, E.A., Wen, J., 2015. Using nuclear gene data for plant phylogenetics: progress and prospects II. Next-gen approaches. J. Syst. Evol. 53, 371–379.