Molecular Phylogenetics and Evolution 181 (2023) 107727

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogenomics of *Aralia* sect. *Aralia* (Araliaceae): Signals of hybridization and insights into its species delimitations and intercontinental biogeography

Jing Liu [a], Ze-Long Nie [b], Chen Ren [c], Chun Su [d], Jun Wen [e,*]

[a] College of Life Science, Sichuan Agricultural University, Ya'an, Sichuan 625014, China
[b] College of Biology and Environmental Sciences, Jishou University, Jishou, Hunan 416000, China
[c] Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, Guangdong 510650, China
[d] Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan 666303, China
[e] Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012, USA

ARTICLE INFO

Keywords:
Aralia sect. *Aralia*
Araliaceae
Phylogenomics
Cytonuclear discordance
Eastern Asian-North American disjunction
Sino-Japanese disjunction

ABSTRACT

Genome-scale data have significantly increased the number of informative characters for phylogenetic analyses and recent studies have also revealed widespread phylogenomic discordance in many plant lineages. *Aralia* sect. *Aralia* is a small plant lineage (14 spp.) of the ginseng family Araliaceae with a disjunct distribution between eastern Asia (11 spp.) and North America (3 spp.). We herein employ sequences of hundreds of nuclear loci and the complete plastomes using targeted sequence capture and genome skimming to reconstruct the phylogenetic and biogeographic history of this section. We detected substantial conflicts among nuclear genes, yet different analytical strategies generated largely congruent topologies from the nuclear data. Significant cytonuclear discordance was detected, especially concerning the positions of the three North American species. The phylogenomic results support two intercontinental disjunctions: (1) *Aralia californica* of western North America is sister to the eastern Asian clade consisting of *A. cordata* and *A. continentalis* in the nuclear tree, and (2) the eastern North American *A. racemosa* forms a clade with *A. bicrenata* from southwestern North America, and the North American *A. racemosa* - *A. bicrenata* clade is then sister to the eastern Asian clade consisting of *A. glabra* (Japan), *A. fargesii* (C China), and *A. apioides* and *A. atropurpurea* (the Hengduan Mountains). *Aralia cordata* is supported to be disjunctly distributed in Japan, Taiwan, the Ulleung island of Korea, and in Central, Southwest and South China, and *Aralia continentalis* is redefined with a narrower distribution in Northeast China, eastern Russia and peninsular Korea.

1. Introduction

Aralia L. sect. *Aralia* is a small group of 14 herbaceous perennial species in the ginseng family Araliaceae, and represents one of the six sections of the genus *Aralia* (Wen, 2011). The section shows a classical disjunct distribution between eastern Asia (11 spp.) and North America (3 spp.), with several closely related species allopartrically distributed in both eastern Asia and North America (Wen et al., 1998; Wen, 2011). Within eastern Asia, species of the section also show strong endemism in each of the major biogeographic regions, such as western Himalaya, eastern Himalaya, Southwest China, Northeast Asia, Japan, and Taiwan, making *Aralia* sect. *Aralia* an ideal system for exploring biogeographic patterns of eastern Asia. *Aralia* represents one of the c. 50 plant genera in Araliaceae (Wen et al., 2001; Frodin & Govaerts, 2003; Lowry et al., 2004; Plunkett et al., 2004, 2018; Gallego-Narbón et al., 2022), with generic delimitation being actively studied (Wen, 1993, 2011; Mitchell et al., 2012; Li & Wen, 2013; Lowry & Plunkett, 2020).

Most species of *Aralia* sect. *Aralia* have been used as traditional medicines in eastern Asia, however, the section has been taxonomically difficult with species delimitations controversial (Wen, 2011). The *Aralia cordata* complex is the largest group in the section, and includes six eastern Asian species: *A. cordata*, *A. taiwaniana*, *A. continentalis*, *A. schmidtii*, *A. tibetana* and *A. cachemirica*, that occur allopatrically throughout eastern Asia and are morphologically similar (Wen, 2011).

* Corresponding author.
E-mail address: wenj@si.edu (J. Wen).

https://doi.org/10.1016/j.ympev.2023.107727
Received 16 July 2022; Received in revised form 20 January 2023; Accepted 3 February 2023
Available online 7 February 2023
1055-7903/Published by Elsevier Inc.

Based on the most recent monograph (Wen, 2011), *Aralia cordata* is recognized to be restricted to Japan and the Ulleung volcanic island near the Korean Peninsula, and *A. continentalis* has a broad distribution from southwest China to eastern Russia and Korea; *Aralia taiwaniana* is endemic to the mountains of central Taiwan; *Aralia schmidtii* is distributed from the Sakhalin and Kuril islands in eastern Russia to the northern part of the Japanese archipelago in the North Pacific; *Aralia tibetana* and *A. cachemirica* are endemic to the eastern Himalaya and the western Himalaya, respectively (Fig. S1). Yet *Aralia continentalis* and *A. taiwaniana* were sometimes treated as synonyms of *A. cordata* or as a variety (Zhu, 1989; Huang, 1993; Xiang and Lowry, 2007). The species delimitations of the *Aralia cordata* complex remain controversial due to the lack of phylogenetic studies with a broad taxon sampling.

Wen et al. (1998) inferred the phylogeny of *Aralia* sect. *Aralia* based on the nuclear ribosomal ITS data. Due to the limited phylogenetic informative sites and limited taxon sampling (*A. glabra*, *A. continentalis*, *A. schmidtii* and *A. taiwaniana* not included), the phylogenetic relationships within the section are still poorly understood. Wen et al. (1998) reported that the three North American taxa, *A. californica*, *A. racemosa*, and *A. racemosa* ssp. *bicrenata* (now recognized as *A. bicrenata* in Wen, 2011) formed a monophyletic group based on ITS sequence data. But *Aralia californica* was sister to the clade consisting of the eastern North American *A. racemosa* and eastern Asian *A. cordata* based on the combined data of ITS and three chloroplast markers (Wen, 2011). The monophyly of the North American group of *Aralia* sect. *Aralia* thus needs to be tested, and resolving the positions of the three North American species will provide an example to understand the evolution of classical eastern Asian-North American biogeographic disjunctions (Wen, 1999; Wen et al., 2016; Lee et al., 2019; Zhou et al., 2020).

The position of the eastern Asian *Aralia henryi* has also been controversial. It was placed outside the core *Aralia* sect. *Aralia* clade with a low support value (BS = 53) in the ITS topology (Wen et al., 1998) while it formed a monophyletic group with the other species of *Aralia* sect. *Aralia* based on the combined data (Wen, 2011). *Aralia henryi* is distributed from southwest to central China, extending to Zhejiang province of eastern China. Morphologically, *A. henryi* bears a small inflorescence with a few umbels and flowers and has bright red fruits, making *A. henryi* highly distinctive in *Aralia*. The phylogenetic position of *A. henryi* hence needs to be further tested.

Genome-scale data including plastid and nuclear genome data have significantly increased the number of informative characters for phylogenetic analyses and have been successfully utilized for reconstructing phylogenies of many plant groups recently (Wen et al., 2013, 2015; Zimmer & Wen, 2015; Xiang et al., 2017; Kleinkopf et al., 2019; Leebens-Mack et al., 2019; Li et al., 2019; Medina et al., 2019; Ma et al., 2021). The chloroplast genome sequences, easily obtainable from genome skimming (Jansen & Ruhlman, 2012), have become a popular source of data for reconstructing the phylogenetic relationships of many groups at different levels in the past ten years (Bock et al., 2014; Panero et al., 2014; Zhang et al., 2015, 2016; Wen et al., 2018; Liu et al., 2019; Valcárcel & Wen, 2019; Wang et al., 2020; Li et al., 2021; Zhou et al., 2022). The nuclear genome harbors biparentally inherited information and is essential to explore evolutionary events, such as hybridization and introgression (Rieseberg & Soltis, 1991; Soltis & Kuzoff, 1995; Ackerfield & Wen, 2003; Hardig et al., 2000; Bock et al., 2014). In the phylogenomic era, considerable discordance across gene trees has been detected in many cases (Salichos & Rokas, 2013; Wen et al., 2013; Xiang et al., 2017; Hodel et al., 2021; Morales-Briones et al., 2018, 2021; Widhelm et al., 2019; Liu et al., 2021a; Ma et al., 2021). The potential sources of conflicts among gene trees may include hidden paralogs, ancient or recent hybridization or reticulate evolution, horizontal gene transfer, and the widespread presence of incomplete lineage sorting (ILS) in recent rapid radiations (Hodel et al., 2022). Furthermore, comparing phylogenies inferred from genomes with different inheritance patterns has the potential to elucidate hybridization and introgression events (Sun et al., 2015; Crowl et al., 2017; Wang et al., 2021; Dong et al., 2022; Zhou et al., 2022), which are often considered as common driving forces in the evolution of plant species (Anderson & Stebbins, 1954; Seehausen, 2004; Soltis & Soltis, 2009; Payseur & Rieseberg, 2016; Liu et al., 2019, 2021a, 2022).

Combining the benefits from the target enrichment technique and the cost-effective genome skimming, the Hyb-Seq approach can efficiently survey hundreds or thousands single- or low-copy nuclear genes with chloroplast genomes as the by-product from the same genomic library, and has been widely employed to resolve phylogenetic relationships in diverse lineages (e.g. Weitemier et al., 2014; Heyduk et al., 2015; Stephens et al., 2015; Moore et al., 2017; Villaverde et al., 2018; Li et al., 2019; Ma et al., 2021). With the utilization of massive genome-scale sequence data for phylogenetic studies, highly supported yet conflicting cytonuclear phylogenies from Hyb-Seq data have become more prominent and need to be carefully assessed (Bruun-Lund et al., 2017; Morales-Briones et al., 2018, 2021; Herrando-Moraira et al., 2019; Edelman et al., 2019; Lee-Yaw et al., 2019; Ma et al., 2023). In addition to the primary goal of constructing species trees, it is also necessary to explore evolutionary processes such as hybridization events that we can infer based on the patterns of discordances of phylogenomic datasets.

In this paper, we employ a phylogenomic dataset of hundreds of nuclear loci and the complete chloroplast genomes through targeted sequence capture and genome skimming to: (1) uncover the phylogenetic relationships within *Aralia* sect. *Aralia*; (2) test and explore the potential discordance among gene trees and cytonuclear discordance; and (3) explore the intercontinental biogeographic disjunction history between eastern Asia and North America in *Aralia* sect. *Aralia*.

2. Materials and methods

2.1. Taxon sampling and library preparations

We sampled 13 of the 14 recognized species (only except *A. schmidtii*) of *Aralia* sect. *Aralia*, with multiple accessions representing each species from across most of their geographic distribution and/or morphological variability, along with six species representing three other sections of *Aralia* and four species of the closely related *Panax* for the phylogenetic analyses (Wen & Zimmer, 1996; Wen & Nowicke, 1999; Zuo et al., 2017). Four taxa of *Polyscias*, *Pseudopanax* and *Osmoxylon* of Araliaceae were used as the outgroups based on the previous phylogenetic studies (Wen et al., 2001; Valcárcel & Wen, 2019). The voucher information is presented in Table 1.

Total genomic DNAs were extracted from silica-gel dried leaves or herbarium specimens using the modified SDS method (Dellaporta et al., 1983). The quality of total DNA was checked by gel electrophoresis and Qubit (dsDNA BR assays: Thermo Fisher Scientific, Waltham, Massachusetts, USA), and then diluted to a concentration of 15 ng/μL. The DNAs were then sheared to a target peak size of ca. 500 bp by sonication in a Q800 Sonicator (Qsonica L.L.C, Newtown, Connecticut, USA) for library preparation. An Illumina paired-end DNA library was constructed using a NEBNext UltraIIDNA Library Prep Kit for Illumina (New England Biolabs). Agencourt AMPure XP magnetic beads (Beckman Coulter) were used for fragment size selection for 400–500 bp, and also for DNA purification.

2.2. Baits, target enrichment and sequencing

The baits used for target enrichment and sequencing were reported in detail in Gallego-Narbón et al. (2022). Briefly, the bait set of targeted genes was designed for Araliaceae based on comparisons of the genomes of *Panax ginseng* (Kim et al., 2018; https://ginsengdb.snu.ac.kr/) and *P. notoginseng* (Zhang et al., 2017; https://www.plantkingdomgdb.com/panax_notoginseng/), and transcriptome sequences of *Hedera helix* and *Polyscias fruticosa* from the 1000 Plant Transcriptomes Initiative (OneKP; https://www.onekp.com; code: SUVN and EDBB, respectively). The selection was conducted mainly in Geneious Prime (https://www.

Table 1
Species sampled in the current study. Voucher specimens are deposited at the United States National Herbarium (US).

Species	Voucher	Geographical origin
Aralia sect. *Aralia*		
Aralia apioides	Tibet-MacArthur 2680	Tibet, China (Qinghai-Tibet Plateau)
Aralia apioides	Tibet-MacArthur 2382	Yunnan, China (Hengduan Mountains)
Aralia apioides	Tibet-MacArthur 2283	Yunnan, China (Hengduan Mountains)
Aralia apioides	J. Liu 809	Sichuan, China (Hengduan Mountains)
Aralia atropurpurea	Tibet-MacArthur 1917	Sichuan, China (Hengduan Mountains)
Aralia atropurpurea	J. Liu 729	Sichuan, China (Hengduan Mountains)
Aralia bicrenata	Wen 1-D1	Colorado, USA
Aralia bicrenata	Wen 1-J4	New Mexico, USA
Aralia cachemirica☆	No. 5964	Changla galli, Pakistan (Himalayas)
Aralia cachemirica	Anzar 501	Kashmir, India
Aralia californica	Wen 6694	Monterey Co., California, USA
Aralia californica	Wen 7104	Tehama Co., California, USA
Aralia californica	Wen 13496	Siskiyou Co., California, USA
Aralia chinensis	Wen 11402	Guangdong, China
Aralia continentalis☆	Russia 1952	Razdolnaya river basin, Russia
Aralia "continentalis"	Wen 12092	Sichuan, China (Hengduan Mountains)
Aralia continentalis	Xie L. 2018	Hebei, China
Aralia continentalis	Wen 11550	Guangxi, China
Aralia continentalis	Wen 3129	Jilin, China
Aralia continentalis	Wen 5426	Jilin, China
Aralia continentalis	Wen 5545	Jiangxi, China
Aralia continentalis	Wen 9330	Hunan, China
Aralia cordata	AJOU 21170	Ullung county, South Korea
Aralia cordata	AJOU 21169	Ullung, South Korea
Aralia cordata	Zhou 2000	Japan
Aralia cordata	Wen 8542	Japan, Honshu
Aralia cordata	Soejima 1072	Japan
Aralia glabra	Soejima 1085	Japan
Aralia henryi	Wen 12841	Hubei, China
Aralia henryi	Wen 10640	Yunnan, China
Aralia racemosa	Xu,Lu& Lutz 114	Vermont, USA
Aralia racemosa	Wen 11752	North Carolina, USA
Aralia racemosa	Wen 11971	Newton Co., Arkansas, USA
Aralia racemosa	Wen 10482	Quebec, Canada
Aralia racemosa	Wen 13427	New Hampshire, USA
Aralia taiwaniana	Wen 9424	Taiwan
Aralia tibetana☆	No. 5606	Painyā Kholā, Nepal (Himalayas)
Aralia tibetana	Q22	Tibet, China (Qinghai-Tibet Plateau)
Aralia tibetana	Xie L. 13145	Tibet, China (Qinghai-Tibet Plateau)
Aralia tibetana	Xie L. 13149	Tibet, China (Qinghai-Tibet Plateau)
Aralia sect. *Dimorphanthus*		
Aralia dasyphylla	Wen 10129	Java, Indonesia
Aralia elata	Liu P.L. 363	Shaanxi, China
Aralia fargesii	Liu P. L. 418	Shaanxi, China
Aralia fargesii	Wen 5446	Hubei, China
Aralia sect. *Humiles*		
Aralia humilis	Wen 8679	Oaxaca, Mexico
Aralia sect. *Pentapanax*		
Aralia franchetii	Wen 14623	Yunnan, China
Aralia parasitica	Wen 14615	Yunnan, China
Panax		
Panax bipinnatifidus	Nie 1162	China
Panax japonicus	Soejima 1118	Japan
Panax quinquefolius	Wen 6243	Virginia, USA
Panax trifolius	Wen 10099	Maryland, USA
Outgroups		
Osmoxylon novoguineense	Wen 10706	Indonesia, Keeron District, Papua
Arthrophyllum javanicum	Wen 10157	Bogor, Indonesia
Polyscias australiam	Wen 10710	Irian Barat, Indonesia
Pseudopanax latevirens	Wen 2019306	Chile originally, cult.

geneious.com), with a maximum E-value of 1×10^{-10} for all BLAST searches. The coding regions of *P. ginseng* (ver. 1.1) and *P. notoginseng* (ver. 1) were BLASTed against each other to identify putative orthologous genes. Only the genes present in both genomes were considered, and we selected the genes that were strictly single-copy in *P. notoginseng* and had no more than two matches in *P. ginseng*, because *P. notoginseng* is a diploid ($2n = 2x = 24$) and and *P. ginseng* is a tetraploid ($2n = 4x = 48$). The selected genes were double checked by BLASTing against their genome data, and then filtered to keep those longer than 300 bp and having 85–99 % of the pairwise identities between *P. ginseng* and *P. notoginseng*. The retained genes were further BLASTed against the transcriptome assemblies of *Hedera helix* and *Polyscias fruticosa*, and those with multiple matches in either *H. helix* or *P. fruticosa* were excluded. This process resulted in a set of 933 genes with a range from 300 to 3990 bp and a total length of 700,778 bp. For bait synthesis, we used the sequences of *Panax ginseng* as the references to check if the genes were single-copy. Otherwise, we used the sequences of *P. notoginseng* or *H. helix*. The myBaits 120-mer biotinylated RNA probes with 3 × tiling on each locus were designed and synthesized by Arbor Biosciences (Ann Arbor, MI, USA, Catalog #180912-94A). Before the synthesis, the probes were filtered by removing high repeat elements, which were identified using the Plant Genome and Systems Biology (PGSB) repeat database (Nussbaumer et al., 2013).

We pooled 6–10 indexed libraries in one reaction, with equimolar amounts of each sample to a total of 100 ng. Solution-based hybridization and enrichment was carried out through the MYBaits target enrichment system (MYcroarray, Ann Arbor, Michigan) following the standard MYBaits v4.01 protocol (https://arborbiosci.com/mybaits-manual/). The final enriched libraries were quantified by gel electrophoresis and Qubit. To recover the plastid regions, 40 % of unenriched libraries were added to the target-enriched libraries, then pooled into one superpool in equimolar ratios. The superpool was then sequenced with 150 bp paired-end reads on an Illumina HiSeq 4000 platform (Novogene, Sacramento, CA, USA).

2.3. Raw data processing and assembly

The raw reads data generated from the Illumina HiSeq runs were trimmed and filtered using Trimmomatic v.0.39 (Bolger et al., 2014), with a 4 bp sliding window, minimum quality scores of 15 and the 36 bp minimum length of reads. The results were checked by FastQC (https://www. bioinformatics.babraham.ac.uk/projects/fastqc) for quality control.

For nuclear gene assembly, the pipeline HybPiper v1.3 (Johnson et al., 2016) was used to extract and sort genes for downstream phylogenetic analyses. In this pipeline BWA (Burrows-Wheeler Alignment tool) (Li & Durbin, 2009) was used to sort reads to target genes, and SPAdes v.3.5.0 (Bankevich et al., 2012) was used to assemble reads to contigs for each target sequence. A target sequence file containing the total 933 genes for bait design was used as reference in the pipeline. If multiple long-length contigs were assembled by SPAdes for a target sequence for a sample, one contig was chosen automatically based on higher sequencing coverage depth or higher percent identity to the reference sequence under the default settings. Then individual locus datasets where each sample was represented by one sequence, was produced and used for subsequent processing. Also paralog assessment was done for all samples and genes.

For plastome assembly, we employed NOVOPlasty (Dierckxsens et al., 2017) for *de novo* assembly based on clean reads that adapters

were removed from raw reads with Trimmomatic. The assembly was prepared with a k-mer of 39, and *rbc*L gene extracted from *Aralia cordata* (Genbank MH778959) as seed sequence. Samples which failed to be assembled by the one step NOVOPlasty, due to likely low sequence coverage of some parts, a multistep approach described by Zhang et al. (2015) and Liu et al. (2019) was used. Briefly, the reference-guided assemblies were performed by Bowtie2 (Langmead & Salzberg, 2012). Then, the chloroplast genomes of *Aralia cordata* (Genbank MH778959), *Panax notoginseng* (Genbank KP036468), and *Osmoxylon novoguineense* (Genbank MK943807) were applied as reference genomes for the first batch assembly performed with SPAdes with K-mer of 21, 33, 55, 77. To correct errors and ambiguities resulted from each approach, the scaffolds from the abovementioned *de novo* assembly were remapped to the plastome based on the reference-guided assembly in Geneious. Subsequently, chloroplast genomes were annotated and manually checked using Geneious, and the abovementioned three reference genomes were used for annotation for the related groups.

2.4. Phylogenetic analyses

2.4.1. Plastome data

For the plastome phylogenetic analyses, the 55 whole plastomes (WPs, including only one copy of inverted repeat region) were aligned using the automatic alignment strategy in MAFFT (Nakamura et al., 2018). Poorly aligned regions were then removed using trimAl v1.2 (Capella-Gutiérrez et al., 2009) with a heuristic selection. To assess the consistency of phylogenetic inference from the whole plastomes (WPs) and the coding sequences only, we extracted a total of 80 coding sequences (CDSs) from each annotated plastome. Sequences were aligned and trimmed with the same options and all resulting alignments were concatenated to a supermatrix in Geneious.

The datasets of WPs and CDSs were used to construct phylogenetic trees using maximum likelihood (ML) and Bayesian inference (BI), respectively. We performed the ML analyses with RA × ML v.8.1.11 (Stamatakis, 2014) using the following settings: GTR + G model, 100 searches for the best-scoring tree, 1000 rapid bootstrap (BS) replicates. For the CDSs dataset, RA × ML analyses were used with partitioned by gene. The BI analysis was performed with MrBayes 3.2.7 (Ronquist et al., 2012). The GTR + G model of molecular evolution was used for WPs dataset, and the best-fit partitioning schemes and models estimated by PartitionFinder2 (Lanfear et al., 2016) were used for the CDSs dataset. Two Markov chain Monte Carlo (MCMC) analyses were run for 10 000 000 generations, respectively. Trees were sampled from every 1000 generations, and the first 25 % of samples were discarded as burn-in. Convergence was assumed when all parameters had effective sampling sizes (ESS) greater than 200 as estimated with Tracer v1.7 (Rambaut et al., 2018). TreeAnnotator v1.10.4 was used to summarize the remaining trees and generate a maximum clade credibility (MCC) tree.

2.4.2. Nuclear phylogenomic analyses

Individual nuclear gene datasets were aligned using MAFFT with the -auto option. Then spurious sequences or poorly aligned regions were automately removed from alignment with options -gt 0.8, -resoverlap 0.75 and -seqoverlap 75 in trimAl. To minimize systematic error and phylogenetic biases, datasets after trimming were then filtered with alignment length greater than 300 bp and with at least 48 samples. Then we obtained a dataset containing 766 loci (nr55s766). Since our study focuses on the phylogenetic relationships of *Aralia* sect. *Aralia*, we reduced our taxon sampling to 10 ingroups of the section with one outgroup (*Panax japonicus*), and constructed three reduced datasets with different screening criteria: nr11s417 (417 nuclear loci having the 11 species), nr11s364 (364 most conserved nuclear loci, based on minimum 80 % identical sites), and nr11s95 (95 nuclear loci with length greater than 1000 bp).

Considering a relatively high proportion of paralogs recovered, we also performed a tree-based orthology inference approach explored by Yang and Smith (2014). The paralog copies for all samples and exons were retrieved through HybPiper. Because it is common that multiple isoforms for each gene may be recovered, they may form monophyletic or paraphyletic tips on the gene tree. For phylogenomic purposes, we performed tip masking to keep the monophyletic tips as the representative on the gene tree, with the highest number of non-ambiguous characters in the alignment. After masking monophyletic tips, paralogy pruning using the MO strategy was conducted to infer orthologs for downstream phylogenetic analyses following the pipeline available in the repository (https://bitbucket.org/yanglab/phylogenomic_dataset_construction/). As a result, a dataset with 500 ortholog loci (nr55s500or) was produced from MO orthology inference. A reduced dataset (nr11s194or) with 11 samples and 194 ortholog loci was also produced through strictly one-to-one paralogy pruning.

For both nuclear datasets (nr55s766 and nr55s500or), we performed concatenation- and coalescent-based methods to estimate phylogenetic species trees. The four reduced datasets (nr11s417, nr11s364, nr11s95 and nr11s194or) were only used to infer coalescent-based phylogenies.

As per each concatenated super-matrix, we inferred a ML tree using RA × ML with a partition-by-locus scheme, a GTR + G model for all partitions, 100 searches for the best tree and 1000 rapid bootstrap replicates for assessing clade support.

To reconstruct the coalescent-based species tree, we employed ASTRAL-III (Zhang et al., 2018), a summary statistic method that accounts for discordance among gene trees produced by ILS. Individual gene trees were estimated using RA × ML with a GTR + G model, and 500 replicates of bootstrapping. As ASTRAL is sensitive to gene tree estimation error (Gatesy & Springer, 2014; Mirarab & Warnow, 2015), we used TreeShrink (Mai & Mirarab, 2018) to remove suspicious tips with abnormally long branched from specific gene trees using a false-positive error rate (α) of 0.01, and collapsed branches with low bootstrap support (<10 %) using the Newick Utilities package (Junier & Zdobnov, 2010). The resulting gene trees were used as input to estimate species trees in ASTRAL-III, and using local posterior probabilities (LPP) to assess clade support.

2.5. Phylogenetic conflict analyses

To explore phylogenetic incongruence between gene tree and species tree estimates, we employed Phyparts (Smith et al., 2015) to calculate internode certainty all (ICA) and the number of conflicting/concordant bipartitions, using the individual gene trees with a BS support cutoff of 50 % and the species tree inferred from ASTRAL as the target tree. Given a set of individual gene trees, ICA value can quantify the degree of conflict on each node of a species tree. ICA values close to 1 indicate strong concordance in the bipartition, ICA values close to 0 indicate equal support for one or more conflicting bipartitions. Negative ICA values indicate internode of interest conflicts with a bipartition that has a higher frequency, and ICA values close to −1 indicate the absence of concordance for the bipartition of interest (Salichos et al., 2014). Phyparts results were visualized with phypartspiecharts.py (available from https://github. com/mossmatters/MJPythonNotebooks).

To assess the amount of gene tree conflict on branches of species tree, we also evaluated tree conflict and branch support with Quartet Sampling (QS; Pease et al. 2018) using 1000 replicates. Quartet Sampling assesses the confidence, consistency, and informativeness of each internal branch through subsampling quartets from the input trees and alignment, with a set of three scores (Quartet Concordance or QC, Quartet Differential or QD, and Quartet Informativeness or QI) (Pease et al., 2018). The QC indicates quartet concordances; QD informs about the proportion of the two possible discordant topologies; and QI notifies the informative capacity of the dataset to resolve the phylogeny. The nuclear dataset (nr55s766) was used for both phylogenetic conflict analyses of Phyparts and QS.

2.6. Phylogenetic network analyses

To explore specific reticulate evolution events and gene flow, we also applied a phylogenetic network analysis using SNaQ in the Julia package PhyloNetworks (Solís-Lemus and Ané, 2016; Solís-Lemus et al., 2017), which uses maximum pseudolikelihood to fit a network and accounts for the contribution of ILS. Considering our focus on the phylogeny of *Aralisa* sect. *Aralia* and the computational limitation, we reduced our taxon sampling to 16 ingroups in the section with *Panax japonica* as outgroup. A dataset of 349 strictly one-to-one ortholog loci with 17 samples (nr17s349or) was produced through paralogy pruning and included in the species networks inference. Networks were inferred based on inheritance probabilities and branch lengths using a pseudolikelihood value, which is a multiple of the network's log-likelihood score.

We tested the maximum number of hybridization events (hmax) with values ranging from 0 to 9, using the ASTRAL tree as a starting tree for the initial network optimization, and 50 independent runs per hmax value. The optimal number of hybridization events was selected by examining the log pseudo-likehood profile of hmax as recommended by the developers (Solís-Lemus & Ané, 2016). For the best value of hmax, the phylogenetic network with the highest pseudo-likelihood values was selected.

2.7. Divergence age estimates

To estimate the disjunction times between eastern Asia and North America, we employed treePL 1.0 (Smith and O'Meara, 2012), an implementation of a divergence time method using penalized likelihood (Sanderson, 2002) for very large datasets.

We first performed an analysis under the ''prime'' option to select the optimal set of parameter values. Then we set gradient-based (opt), auto-differentiation-based (optad) and auto-differentiation cross-validation-based optimizers (optcvad) to two and ran a second analysis using random subsample and replicate crossvalidation to identify the best value for the smoothing parameter. We ran the final analysis by setting the best chi-square value for smoothing parameter (smoothing = 0.1). To compare the results of different datasets, ML tree inferred from the concatenated nuclear matrix (nr55s766) and plastid matrix (WPs) were used as input tree, respectively.

For both date estimations, two secondary calibration points and one fossil calibration point were set. We constrained the *Aralia* - *Panax* divergence at 34.44 Mya (95 % highest posterior density or HPD:20.22–50.55 Mya), and the root of the tree was calibrated to be 52.67 Mya (95 % HPD: 33.21–74.66 Mya) based on two secondary calibration points obtained from the complete plastome matrix in Valcárcel & Wen (2019). The crown age of *A. elata* - *A. dasyphylla* - *A. chinensis* clade was constrained to be 11.2 Mya based on fossil seeds of *Aralia* from Stare Gliwice in Upper Silesia of eastern Europe in the Miocene, which were closely related to the extant *Aralia elata* (Szafer 1961). For each date estimations, we ran three separate analyses with the mean age, the minimum age and maximum age inferred in Valcárcel & Wen (2019), to estimate the range of the node age.

3. Results

3.1. Characteristics of the datasets

For all 55 samples, percentage of target reads out of total number of reads ranged from 6.1 % (*A. continentalis* 1952) to 65.7 % (*A. cordata Wen 5426*); assembly through HybPiper resulted in a total of 933 nuclear genes (100 % of targeted sequences) and up to 931 genes per species; and paralogous copies were identified for a mean of 99 genes per species (from 27 to 247 genes), about 10.6 % of genes appear to have paralogs recovered. Statistics on gene assembly are presented in Table S1.

A total of 766 low-copy nuclear loci (at least 300 bp alignment length and 48 sample coverage) were obtained from the original 933-locus datasets assembled by HybPiper. After paralog pruning, a dataset of 500 ortholog loci were obtained. The two datasets were used for phylogenetic inference. The aligned length of the concatenated 766-locus matrix was 586,741 bp, of which 73,797 characters (12.6 %) were parsimony-informative, while the 500-locus matrix was 422,677 bp in length, of which 47,900 characters (11.3 %) were parsimony-informative.

We also successfully assembled the complete plastomes of all 55 samples from off-target reads, of which 43 were assembled by NOVOPlasty. The plastome matrix (WPs) had an aligned length of 130,320 bp with 3925 parsimony-informative sites (3 %). Eighty protein coding regions were retrieved from the each of the annotated plastomes, resulting in a concatenation matrix (CDSs) of 68,526 bp with 1489 parsimony-informative sites (2.2 %).

3.2. Nuclear phylogenomic analyses

Coalescent-based analyses based on the nr55s766 or nr55s500or datasets produced almost identical well-supported topologies on the main clades only with slight differences among shallow nodes at the individual level (Figs. 1 & S2). For both nr55s766 and nr55s500or datasets, the ML tree inferred by concatenated analyses was similar to the coalescent-based ASTRAL tree, only with the exception of the position of *A. taiwaniana* within the *A. cordata* – *A. continentalis* clade and *A.* "*continentalis*" (*Wen 12092*) within the *A. cachemirica* – *A. tibetana* clade (Figs. 1, 2A, S2). Topologies of species trees inferred from the four reduced datasets (nr11s95, nr11s194or, nr11s364 and nr11s417) were identical (Fig. S3) and congruent with the results based on the nr55s766 and nr55s500or (Figs. 1 & S2). Here we present the nuclear phylogeny with ASTRAL species tree based on the nr55s766 dataset (Fig. 1), considering the coalescent-based method is more appropriate for the high heterogeneity of our nuclear dataset.

The monophyly of *Aralia* sect. *Aralia* was supported. Six well-supported clades were identified within the section (Fig. 1). *Aralia henryi* represented the first clade and it was sister to the remaining species of sect. *Aralia*. The three North American species did not form a clade. The southwest North American *A. bicrenata* and the eastern North American *A. racemosa* were sisters (LPP = 1, BS = 100), then the *A. bicrenata* – *A. racemosa* clade was sister to a clade consisting of *A. glabra*, *A. atropurpurea*, *A. fargesii*, and *A. apiodies* from eastern Asia, with strong support (LPP = 1, BS = 100), while the western North American *A. californica* grouped with the eastern Asian *A. cordata* – *A. continentalis* group with moderate support (LPP = 0.97, BS = 90).

Within the *A. cordata* - *A. continentalis* clade, two subclades were recognized: one subclade included samples of *A. continentalis* from Northeast China and Far East Russia, and the second subclade had samples of *A. continentalis* from Central, East and South China, *A. cordata*, and *A. taiwaniana*. Hence *Aralia continentalis* was resolved to be non-monophyletic. In the coalescent-based species tree, *A. taiwaniana* grouped with *A. continentalis* from Central, East and South China with a moderate support value (Fig. 1, LPP = 0.96) and then sister to *A. cordata* from Japan and the Ullung Island, while in the concatenated trees, it was nested within the clade containing samples of *A. continentalis* from Central, East and South China, and *A. cordata* from Japan and the Ulleung Island (Fig. 2A).

Although most of the major clades of *Aralia* sect. *Aralia* and relationships within and among them were similar and well-support in all nuclear topologies (Figs. 1, 2A, S2) based on concatenated and coalescent-based method, a prevalence and high level of discordance between gene trees and species tree were revealed by the ICA scores and QS values at both backbone and shallow nodes (Fig. 1). The sister relationship between the *A. racemosa* + *A. bicrenata* clade and the *A. glabra* – *A. fargesii* group was supported by only two gene trees (ICA = -0.01), and the sister relationship of *A. californica* and the *A. cordata* –*A. continentalis* group was supported by only three gene trees (ICA =

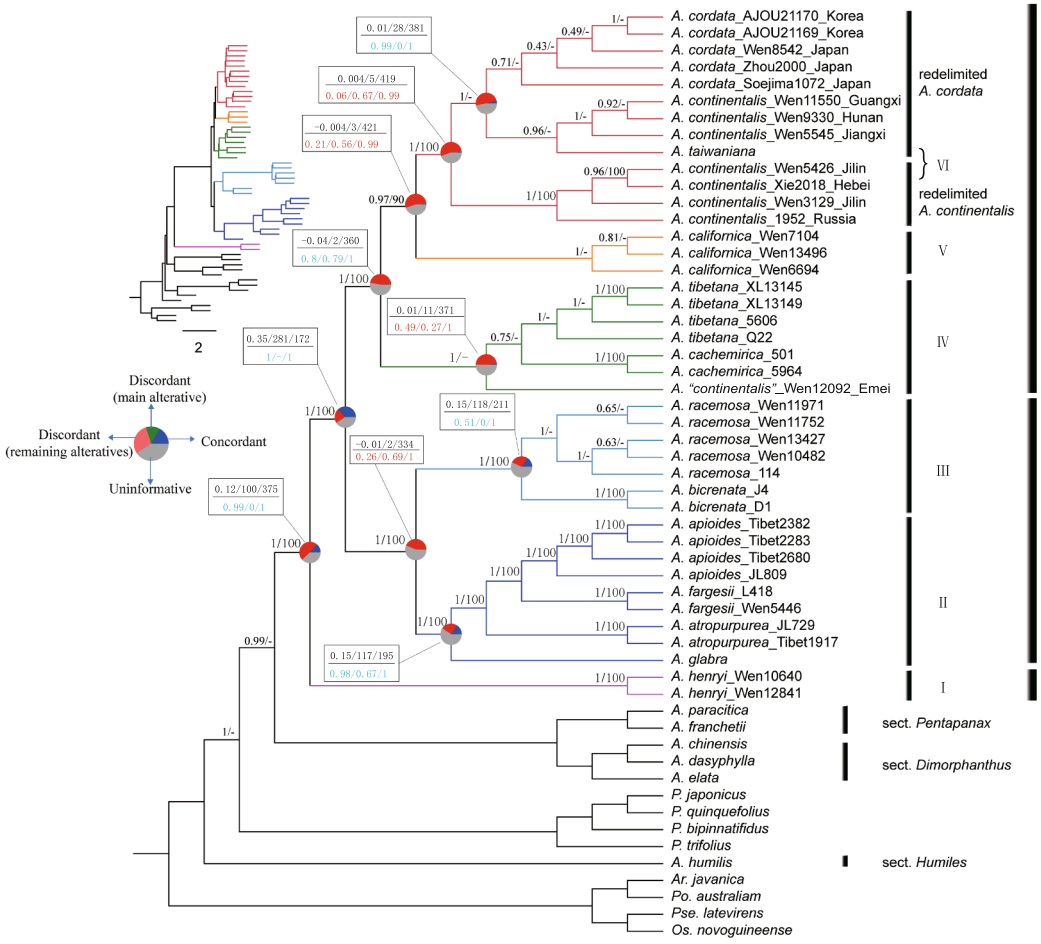

Fig. 1. ASTRAL species tree of *Aralia* sect. *Aralia* inferred from the nr55s766 dataset. Support values (Bootstrap/Local posterior probability) are plotted above branches. For major clades, pie charts inferred from Phyparts present the proportion of gene trees that support that clade (blue), the proportion that support the main alternative bifurcation (green), the proportion that support the remaining alternatives (red), and the proportion (conflict or support) that have < 50 % bootstrap support (gray). Boxes contain gene tree conflict and Quartet Sampling (QS) scores for major clades. In each box, numbers on the upper indicate the Internode Certainty All (ICA) score/the number of gene trees concordant/conflicting with that node in the species tree, and numbers on the lower indicate QS scores: Quartet Concordance/Quartet Differential/Quartet Informativeness with that node in the species tree. QS scores in blue indicate a strong majority of quartets support the focal branch, while scores in red indicate a weak majority of quartets support the focal branch. All node concordance scores shown in Supplementary Fig. S6. Inset shows species tree with branch lengths. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

-0.04). Gene counts (pie charts) based on conflict analyses showed most genes were uninformative for resolving the relationships among the major clades and even among and within species of sect. *Aralia*.

According to the SNaQ analyses, the log-pseudolikelihood value decreased sharply when the hmax from 0 to 1, stabilized at hmax = 2, and improved slowly afterwards (Fig. 3A). Therefore, hmax = 2 was considered as the optimal number of hybridization events, and the phylogenetic network was presented in Fig. 3B. The phylogenetic network suggested two gene flow or hybridization events: *A. racemosa* having a hybridization edge with *A. glabra*, and *A. taiwaniana* having a hybridization edge with *A.* "*continentalis*" (*Wen 12092*). The first gene flow or hybridization event contributed to the origin of *A. racemosa* with inheritance probabilities of 38.9 % from *A. glabra*. The second reticulation event was inferred between *A. taiwaniana* and *A.* "*continentalis*" (*Wen 12092*), with *A.* "*continentalis*" (*Wen 12092*) contributing 10.4 % gene flows to *A. taiwaniana* (Fig. 3B).

3.3. Plastome phylogeny and cytonuclear incongruence

The plastome ML trees inferred from WPs and BI tree inferred from WPs and CDSs data are highly consistent (Fig. 2B & S5). Two samples of sect. *Pentapanax* were nested in the sect. *Aralia*, that resulted in non-monophyly of sect. *Aralia*, probably due to incomplete sampling at the sectional level in *Aralia* (lack of sampling in sect. *Nanae* and sect. *Sciadodendron*). Within *Aralia* sect. *Aralia*, the same five main clades (except the clade (*A. racemose*, *A. bicrenata*)) were recovered. Within each main clade, relationships were mostly congruent with our nuclear analyses. However, the relationships among the major clades differed from the nuclear tree, mainly due to the placement of the three North American species. In the plastome tree, *Aralia racemosa* and *A. bicrenata* formed successive sisters to the *A. cordata* - *A. continentalis* group. *Aralia californica* was sister to the clade containing the *A. cordata* complex, *A. racemosa* and *A. bicrenata* (Fig. 2B).

The position of *A. taiwaniana* was incongruent between the WPs tree and CDSs tree (Fig. 2B & S4). *Aralia taiwaniana* was sister to *A. cordata* from Japan and Ulleung in the WPs tree (Fig. 2B & S4A), while in the CDSs tree it formed a sister group with the clade containing *A. cordata* from Japan and Ulleung and *A. continentalis* from Central, East and South China (Fig. S4B). The relationships within *A.fargesii* - *A.apioides* - *A. atropurpurea* clade were incongruent between the WPs tree and CDSs tree. In the WPs tree, *A.fargesii* was sister to *A.apioides* while in the CDSs tree it formed a sister group with *A. atropurpurea* (Fig. S4).

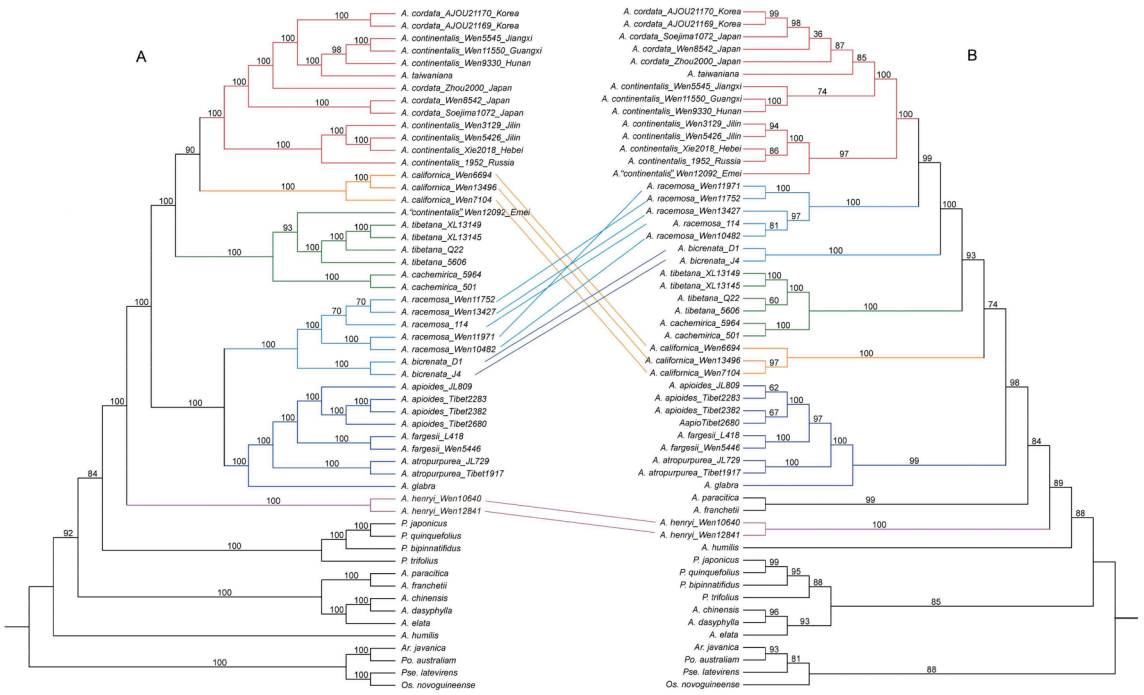

Fig. 2. Phylogenetic trees of *Aralia* sect. *Aralia* with the tanglegram to compare the nuclear and plastome trees. A. The maximum likelihood tree of the concatenated nuclear dataset (nr55s766). B. The maximum likelihood tree of the whole plastome data. Bootstrap values are plotted above branches.

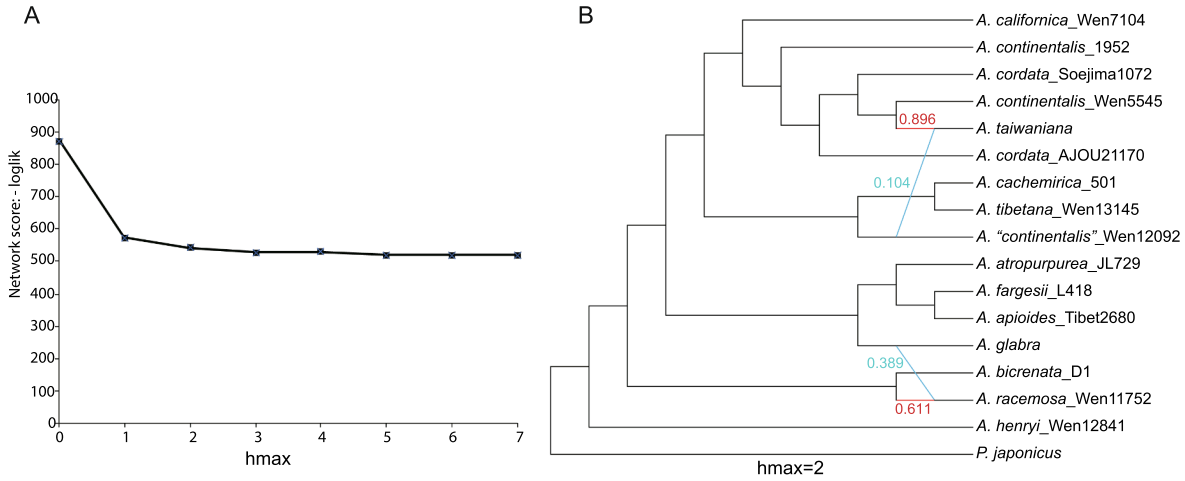

Fig. 3. SNaQ log pseudo-likelihood values of hmax from 0 to 7 (A). Species networks of *Aralia* sect. *Aralia* inferred from SNaQ based on the nr17s349or dataset (B). Red and blue indicates the major and minor edges, respectively, of hybrid nodes. Number next to the branches indicates inheritance probabilities for each hybrid node. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Additionally, the sample of *Aralia* "*continentalis*" (*Wen 12092*, collected from Mt. Emei in Sichuan) grouped with samples of *A. continentalis* from northeastern China in the plastome trees (Fig. 2B & S4), quite different from its placement in the nuclear tree, where it grouped with the two Himalayan species, *A. tibetana* and *A. cachemirica* (Figs. 1 & 2A).

3.4. Divergence age estimates

The origin of sect. *Aralia* was estimated to the late Eocene (38.15 Mya, ranging from 22.15 to 50.02 Mya) (Fig. 4) based on the nuclear data. The three North American species diverged from their eastern Asian congeners during the late Oligocene. The western North American *A. californica* diverged at 23.17 Mya (ranging from 14.44 to 32.94 Mya), and the most recent common ancestors (MRCA) of the North American *A. racemosa* and *A. bicrenata* diverged at 25.32 Mya (ranging from 15.68 to 36.1 Mya). The split of the two subclades within the *A. cordata* - *A. continentalis* group was dated to 21.59 Mya (ranging from 13.47 to 30.67 Mya). Divergence times estimated based on the plastome data were younger (Fig. S5), even though all dating analyses were inferred with the same calibration points. This difference of molecular dating may be caused by the differences in the inheritance pattern and evolutionary rates between plastome and nuclear genome, or also due to the high heterogeneity of the nuclear genes. Plastomes are known to be highly conserved, with relatively low nucleotide substitution rates compared to the nuclear genome (Wolfe et al., 1987).

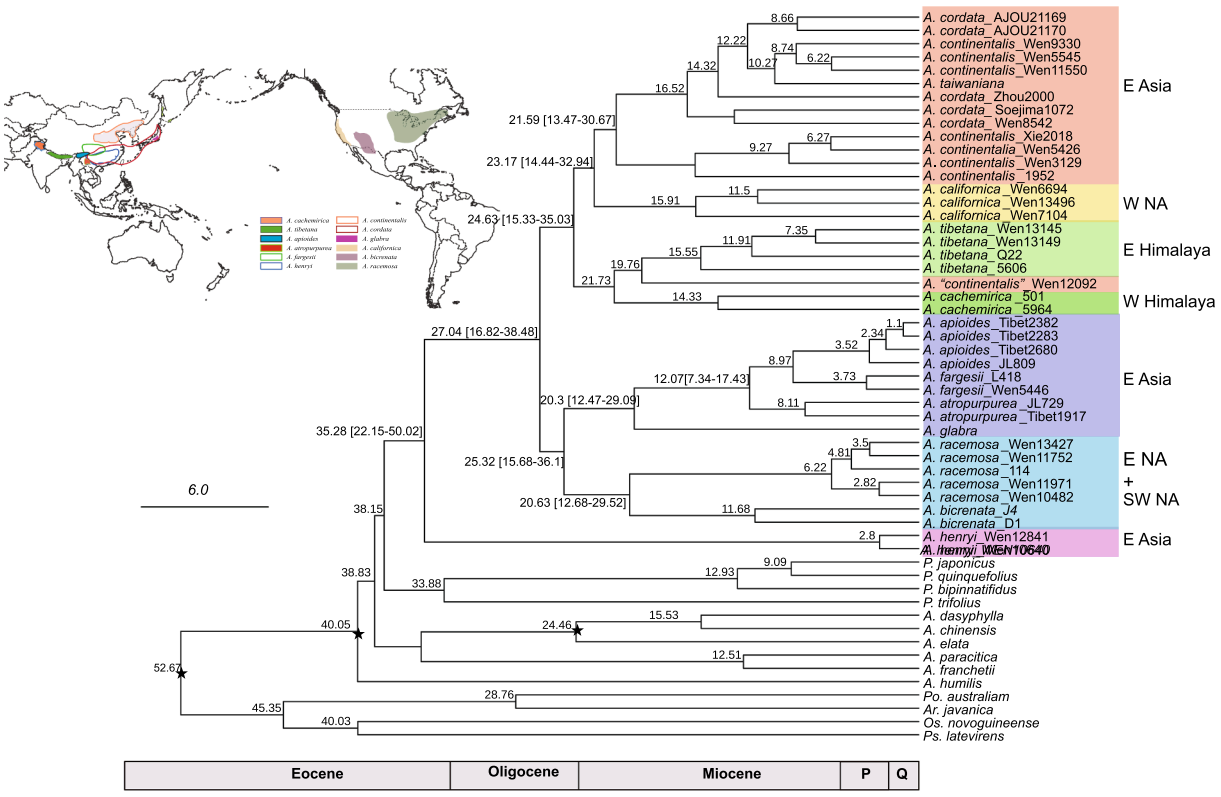

Fig. 4. Divergence time estimates of *Aralia* sect. *Aralia* using treePL based on the nr55s766 dataset. Numbers near nodes shown the estimated node age in millions of years (Mya), a range of node age presented in brackets for focal nodes. Asterisk on nodes represent the calibration points used in dating analysis. In the scale axis, "P" and "Q" correspond to Pliocene and Quaternary, respectively.

4. Discussion

4.1. Cytonuclear discordance

Although most main clades within *Aralia* sect. *Aralia* were recovered by both nuclear and plastome trees, significant cytonuclear incongruence was detected especially related to the positions of the three North American species (Figs. 1 & 2). The nuclear phylogeny supported that *A. californica* was sister to *A. cordata* - *A. continentalis* (LPP = 0.97, BS = 90) with weak QS support (0.21/0.56/0.99) and only 3 out of the 424 informative gene trees (ICA = −0.004) (Fig. 1), while the plastome trees placed *A. californica* as sister to a large clade containing *A. tibetana* - *A. cachemirica*, *A. bicrenata*, *A. racemosa*, and *A. cordata* - *A. continentalis* albeit with low support (BS = 74, Fig. 2B). In the nuclear phylogeny, the *A. racemosa* - *A. bicrenata* clade was sister to the clade consisting of four Asian species (*A. apioides*, *A. fargesii*, *A. atropurpurea*, and *A. glabra*) (LPP = 1, BS = 100) with weak QS support (0.26/0.69/1) and only 2 of the 336 informative gene trees (ICA = −0.01) (Fig. 1), while in the plastome tree, *A. racemosa* and *A. bicrenata* formed successive sisters to the *A. cordata* - *A. continentalis* group (Fig. 2B).

Discrepancies among the phylogenies of nuclear and plastome data can be caused by phylogenetic uncertainties, ILS, hybridization and introgression (Rieseberg & Soltis, 1991; Maddison, 1997; Huelsenbeck et al., 2000; Vargas et al., 2017). Hybridization and introgression events had been elucidated by comparing the phylogenies inferred from nuclear and chloroplast data in many cases (Bock et al., 2014; Sun et al., 2015; Vargas et al., 2017; Rose et al., 2021). In our case, the conflicts are mainly located in the phylogenetic positions of the three North American species between the plastome (WPs & CDSs, Fig. 2B & S4) and nuclear (Figs. 1 & 2A) trees, and also there is low gene tree concordance related to their nodes. The SNaQ results suggested that gene flow or hybridization events in the evolutionary history of the section, with at least two exchanges of genetic material that have occurred between unrelated lineages (Fig. 3B). The cytonuclear discordance may be explained by the likely hybrid origin of some of the species. However, ancestral hybridizations or extinction of the ancestral lineages may obscure the phylogenetic signal among extant species. The hybridization events and processes within *Aralia* sect. *Aralia* need to be further explored in future studies especially using phylogeographic analyses.

4.2. Species delimitations of *A. cordata* and *A. continentalis* and a potential new species from Sichuan province

The species delimitation in the eastern Asian *Aralia cordata* complex has been difficult and controversial (Wen, 2011). Kitagawa (1935) separated *Aralia continentalis* from the Japanese *Aralia cordata*. He recognized *A. continentalis* as from Manchuria of northeastern China as well as Korea, based on the main differences in overall inflorescence architecture, pedicel length, thickness and pubescence, and the connation of styles. Hoo and Tseng (1978) recognized *A. continentalis* as widely distributed from continental eastern Asia, including Northeast, North and Southwest China, as well as Russia and Korea, and *A. cordata* as from East and Central China, Taiwan, and also in Japan. This circumscription of both species was followed by Xiang and Lowry (2007). Hoo and Tseng (1978) also pointed out that inflorescences and pedicels cannot be used as the criteria for species delimitation in *Aralia* because these characters are highly variable even within individuals, however the connation of styles may be a main difference of the two species, with *A. continentalis* having styles connate to middle, and *A. cordata* possessing styles divided nearly to the base. Wen (2011) recognized *A. continentalis* as broadly distributed from Far East Russia, the Korean Peninsula, and Northeast to Central, Southwest and South China, similar to the treatment of Hoo & Tseng (1978) and Xiang and Lowry (2007), while Wen (2011) treated *A. cordata* as restricted to Japan and the Ulleung Island of Korea.

Aralia continentalis as previously delimited is inferred to be non-

monophyletic by our phylogenetic analyses (Figs. 1 & 2). Samples from Northeast and North China and Far East Russia clearly formed a clade with strong support, while samples from Central, East and South China (Guangxi, Hunan and Jiangxi provinces) constituted another clade. *Aralia taiwaniana* from Taiwan are nested within *A. cordata* in the nuclear trees (Figs. 1 & 2A) or sister to *A. cordata* from Japan and the Ulleung Island of Korea in the chloroplast phylogeny (Fig. 2B). Our results thus support defining *A. cordata* more broadly to include populations from Japan, the Ulleung Island of Korea, Taiwan, and Central, East, South and Southwest China. Furthermore, the results support delimiting *A. continentalis* more narrowly as from Northeast China and Far East Russia and the Korea Peninsula. Distributions of the redelimited *A. cordata* and *A. continentalis* are shown in Fig. 1.

Aralia cordata, as it is redefined herein, is characterized by its longer main inflorescences, terminal umbels mostly with more than 35 flowers, and longer pedicels (9–20 mm). The redefined *Aralia continentalis* is characterized by its shorter inflorescences, terminal umbels with<35 flowers, and relatively short (6–10 mm) pedicels.

The geographic separation of *A. continentalis* in Northeast Asia and *A. cordata* in Japan and East, Central and South China supports a north–south biogeographic barrier within the Sino-Japanese Floristic region (also see Wang et al., 2016; Ma et al., 2018a, 2018b, 2022). A climatic barrier that occurred in the early Miocene may have led to such a north–south division (Guo et al., 2008). In the mid to late Miocene, the arid belt expanded and it almost reached the coast of northern China by the Pliocene, serving as a climate barrier that impeded gene flow between the Neogene relict plants in the north and south regions in eastern Asia. Phylogeographic analyses of these two closely related taxa in *Aralia* sect. *Aralia* need to be conducted to clarify the temporal and spatial biogeographic diversification patterns and processes in eastern Asia, as this biogeographic pattern remains emerging and needs to be further studied.

With regard to the *Aralia* "*continentalis*" sample *Wen 12092* from Mt. Emei, the nuclear and plastome data of the population suggest that it may represent a new species from Sichuan province. This sample groups with two Himalayan species *A. tibetana* and *A. cachemirica* in the nuclear tree (Figs. 1 & 2A), while it forms a clade with *A. continentalis* in the plastome tree (Fig. 2B). Phylogenetic network suggested *A.* "*continentalis*" (*Wen 12092*) contributed 10.4 % of gene flows to *A. taiwaniana*. Though the hybridization events detected with SNaQ are not fully congruent with the hybridization events suggested by the conflicts detected between genomes, species network implied a reticulate evolution event involved in this sample.

Morphologically *Wen 12092* from Mt. Emei shows similarities with *A. tibetana* with fewer flowers and non-leafy stipules. But its grouping with *A. continentalis* from North and Northeast Asia in the plastid tree suggests the need to further explore the evolution of this species in the phylogeographic context of *Aralia* sect. *Aralia*. Wen (2011) also noted a few specimens from Sichuan which are similar to *A. tibetana.* Further investigations with more collections around Southwest China are required to explore the role of the hybridization in the origin of this potential new species.

4.3. Phylogenetic relationships within Aralia sect. Aralia

Although significant gene tree discordance and cytonuclear conflict were detected in *Aralia* sect. *Aralia*, all nuclear phylogenies estimated from different datasets and approaches (e.g., ASTRAL and ML trees based on nr55s766, Figs. 1 & 2A; ASTRAL and ML trees based on nr55s500, Fig. S2; ASTRAL trees based on the four reduced nuclear datasets, Fig. S3) resulted in similar topologies: *Aralia* sect. *Aralia* is supported to be monophyletic, and six distinct major genetic lineages were detected within the section (Figs. 1 & 2A). Our discussions here on the phylogenetic relationships of the section are based on the coalescent-based species tree inferred from nuclear data (Fig. 1).

The lineage I consists of only the morphologically unique *A. henryi*. The phylogenetic placement of *A. henryi* has been difficult to resolve (Wen et al., 1998; Wen, 2011). It has been treated as a member of *Aralia* sect. *Anomale* (now placed in synonymy of *Aralia* sect. *Aralia* in Wen, 2011) along with *A. apioides*, *A. fargesii* and *A. glabra* (Harms, 1897; Hoo & Tseng, 1978). Here *A. henryi* represents the first diverged clade of sect. *Aralia*, in agreement with the results from ITS and chloroplast markers (Wen et al., 1998; Plunkett et al., 2004; Wen 2011). *Aralia henryi* is hence best treated as a member of sect. *Aralia* in spite of its morphological distinctions with bright red fruits and highly reduced inflorescence.

The lineage II consists of four eastern Asian species: *A. glabra*, *A. atropurpurea*, *A. apioides* and *A. fargesii*. This clade of the four eastern Asian species is supported in both the nuclear and the chloroplast trees, with *A. glabra* from Japan sister to the clade of the three remaining taxa of this clade from China (Figs. 1, 2A, S4). The four species are morphological similar in terms of leaf architecture and inflorescence structure. Within this group, *A. apioides* from the Hengduan Mountains (HDM) is more closely related to *A. fargesii* from Central China, rather than to another HDM species *A. atropurpurea*. Biogeographically, *A. atropurpurea*, *A. apioides* and *A. fargesii* are distributed in HDM or adjacent to this region, which as a well-known temperate biodiversity hotspot (Xing & Ree, 2017). The origin of the three HDM species was herein estimated in the middle Miocene (12.07 Mya: 7.34–17.43 Mya) (Fig. 4). The rapid and recent uplifts of the HDM during the late Miocene have been an important mechanism driving the diversification and speciation in this region (Wen et al., 2014; Xing & Ree, 2017; Liu et al., 2021b; Mao et al., 2021).

The lineage III consists of two North American species, *A. bicrenata* and *A. racemosa*. A close relationship of these two taxa was suggested based on the ITS sequences (Wen et al., 1998). *Aralia bicrenata* was once treated as a synonym of *A. racemosa* (Smith, 1944) or as a subspecies of *A. racemosa* (Welsh & Atwood, 1975). *Aralia racemosa* and *A. cordata* were suggested as a species pair disjunctly distributed between E Asia and E North America (Li, 1952, 1972). The *A. racemosa* - *A. bicrenata* clade was herein shown to be sister to the clade consisting of four Asian species: *A. apioides*, *A. fargesii*, *A. atropurpurea*, and *A. glabra* in the nuclear tree. Yet in the plastome tree, *A. bicrenata* and *A. racemosa* did not form a clade and these two species were grouped with the *A. cordata* - *A. continentalis* group from eastern Asia (see Fig. 2B).

The two Himalayan species *A. cachemirica* and *A. tibetana* constitute the lineage IV. The western Himalayan *A. cachemirica* and the eastern Himalayan *A. tibetana* are sister species and their divergence was perhaps driven by allopatric speciation in the Himalayan region. Based on morphological similarity and biogeographic affinity, the two species were included in the *Aralia cordata* complex along with *A. cordata*, *A. continentalis*, *A. taiwaniana* and *A. schmidtii* (Wen, 2011). Within eastern Asia, these two Himalayan species indeed showed a close relationship with the Sino-Japanese *A. cordata* - *A. continentalis* group (Figs. 1 & 2A).

The lineage V consists of only the western North American *A. californica. Aralia californica* was sister to the eastern Asian *A. cordata* - *A. continentalis* clade, rather than with the two other North American species. This relationship is consistent with the close biogeographic connection between eastern Asia and western North America, which has been shown in many lineages of vascular plants (Wen et al., 2016, 2018).

The lineage VI includes the two eastern Asian taxa *A. cordata* and *A. continentalis*. Within the redefined *A. cordata*, two biogeographic lineages were suggested: one from Japan and Ulleung island, and the other from continental China and Taiwan. The floristic relationship of eastern China with Japan was considered to be closer than that between eastern China and northeast China and Korea (Xie, 1997). A close relationship between populations from South Central China and Taiwan inferred by all nuclear analyses suggested that populations of *A. cordata* from Taiwan most likely derived from *A. cordata* of mainland China instead of from Japan. *Aralia cordata* is distributed in Fujian province as its eastern limit in continental China, and the Taiwan strait may have

played an important role in the populational exchanges of this lineage.

Of all the six lineages, the *A. glabra* - *A. fargesii* group and the *A. cordata* complex both showed a biogeographic disjunction between Japan and China, a common biogeographic pattern within eastern Asia (Qiu et al., 2011; Qi et al., 2012, 2014; Wang et al., 2020). These two disjunctions in *Aralia* sect. *Aralia* were dated to the early or middle Miocene (Fig. 4). Japan was separated from the mainland East Asia during the Miocene by the opening of the Japan Sea around 15–23 Mya (Otofuji et al., 1985). The close biogeographic relationships between Japan and eastern and Central China have been well documented (Grisebach, 1872; Diels, 1901; Wu & Wu, 1996; Wang et al., 2020).

4.4. Two intercontinental disjunctions between eastern Asia and North America in Aralia sect. Aralia

Aralia sect. *Aralia* is an example showing an uneven intercontinental disjunction between eastern Asia and North America with a noticeably higher species richness in eastern Asia (ca. 10 of the 13 species in the section). Eastern Asia was also suggested as the center of diversity of this section (Wen et al., 1998; Wen, 2011). Biogeographic analyses based on the ITS sequences suggested that the origin of the section in eastern Asia, and the ancestor of the North American species migrated across the Bering land bridge (BLB) in the late Tertiary and diversified within the North America most likely in the west-east direction (Wen et al., 1998). Our phylogenetic analyses based on both nuclear and plastome data suggested that the three North American species may have derived from different ancestral lineages of eastern Asia, and the SNaQ network analyses suggested that eastern Asian species *A. glabra* was involved in the evolutionary origin of the North American species *A. racemosa* through gene flow or hybridization events. The current disjunction pattern may be due to two independent migration events. Although close phylogenetic and biogeographic relationships between eastern and western North America were indicated in many cases (Xiang et al., 1998), even in the previous study of sect. *Aralia* (Wen et al., 1998), there are also some exceptional examples showing that the eastern and western North American elements had derived from different lineages independently (Fritsch, 2001).

The two eastern Asian - North American disjunction events were estimated from the late Oligocene to the early Miocene. It suggested that the Bering Land Bridge (BLB) was the most likely corridor for migration and subsequent exchanges of plants within this temperate section. During the Oligocene to Miocene, BLB was suitable for the temperate deciduous plants (Wen et al., 2016). Our estimate of the divergence of the southwestern North American *A. bicrenata* and eastern North American *A. racemosa* was at around 20.63 Mya (ranging from 12.68 to 29.52 Mya). The Rocky Mountains which uplifted in the early Tertiary in western North America had served as an effective barrier for floristic exchanges between eastern and western North America (Graham, 1993; also see Wen, 1999; Wen et al., 2016).

5. Conclusions

The targeted sequence capture is effective to clarify the phylogenetic relationships of *Aralia* sect. *Aralia*, especially to resolve the relationships within the eastern Asian *Aralia cordata* complex. Redelimitations of *A. cordata* and *A. continentalis* are herein proposed based on the phylogenomic framework, and *Aralia taiwaniana* is best treated as a synonym of *Aralia cordata*. The phylogenomic analyses resolved two intercontinental disjunctions between eastern Asia and North America within *Aralia* sect. *Aralia*: one between eastern Asian and eastern-southwestern North America, and the other between eastern Asia and the Californian Floristic Province of western North America, suggesting eastern and western North American lineages had different origins and migrated along different routes (also see Wen et al., 2016). Phylogenomic data of *Aralia* sect. *Aralia* also showed discordance between gene trees and species trees, as well as extensive cytonuclear discordance. The analyses presented here support hybridization as an important driver in the diversification of *Aralia* sect. *Aralia* both in eastern Asia and North America.

CRediT authorship contribution statement

Jing Liu: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft. **Ze-Long Nie:** Software, Formal analysis. **Chen Ren:** Software, Formal analysis. **Chun Su:** Investigation. **Jun Wen:** Conceptualization, Writing – review & editing, Resources, Supervision.

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.

Data availability

Data will be made available on request.

Acknowledgement

All computational analyses were conducted on the Smithsonian Institution High Performance Computing Cluster (SI/HPC, “Hydra”; https://doi.org/10.25572/SIHPC). We thank Gabriel Johnson for his help with the target enrichment experiment, the United States National Herbarium for collection access, and two reviewers for their constructive suggestions and great interest in our work. We acknowledge the staff in the Laboratories of Analytical Biology at the National Museum of Natural History, the Smithsonian Institution for technical support and assistance. This study was supported by the Smithsonian Barcode Network and the China Scholarship Council under Grant [201806915015].

Data Availability Statement

The raw sequence data (FASTQ files) were deposited in the NCBI Sequence Read Archive (SRA) database under the BioProject PRJNA841627. Alignments and gene trees of all datasets used in this study are available from the Dryad Digital Repository (https://doi.org/10.5061/dryad.tqjq2bw3f).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107727.

References

Ackerfield, J., Wen, J., 2003. Evolution of *Hedera* (the ivy genus, Araliaceae): Insights from chloroplast DNA data. Int. J. Plant Sci. 164, 593–602.
Anderson, E., Stebbins, J.R.G.L., 1954. Hybridization as an evolutionary stimulus. Evolution 8, 378–388.
Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., 2012. Spades: A new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.
Bock, D.G., Kane, N.C., Ebert, D.P., Rieseberg, L.H., 2014. Genome skimming reveals the origin of the jerusalem artichoke tuber crop species: Neither from Jerusalem nor an artichoke. New Phytol. 201, 1021–1030.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for illumina sequence data. Bioinformatics 30, 2114–2120.
Bruun-Lund, S., Clement, W.L., Kjellberg, F., Rønsted, N., 2017. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). Mol. Phylogenet. Evol. 109, 93–104.
Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. Trimal: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Crowl, A.A., Myers, C., Cellinese, N., 2017. Embracing discordance: Phylogenomic analyses provide evidence for allopolyploidy leading to cryptic diversity in a mediterranean *Campanula* (Campanulaceae) clade. Evolution 71, 913–922.
Dellaporta, S.L., Wood, J., Hicks, J.B., 1983. A plant DNA minipreparation: Version ii. Plant Mol. Biol. Report. 1, 19–21.
Diels L. 1901. Die Flora von Central-China. Leipzig: W. Engelmann.Dierckxsens N, Mardulyn P, Smits G. 2017. Novoplasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* 45: e18.
Dierckxsens, N., Mardulyn, P., Smits, G., 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 28, 45(4):e18.
Dong, W.P., Sun, J.H., Liu, Y.L., Xu, C., Wang, Y.H., Suo, Z.L., Zhou, S.L., Zhang, Z.X., Wen, J., 2022. Phylogenomic relationships and species identification of the olive genus *Olea* (Oleaceae). J. Syst. Evol. 60, 1263–1280.
Edelman, N.B., Frandsen, P.B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R.B., Garcia-Accinelli, G., Van Belleghem, S.M., Patterson, N., Neafsey, D.E., Challis, R., Kumar, S., Moreira, G.R.P., Salazar, C., Chouteau, M., Counterman, B.A., Papa, R., Blaxter, M., Reed, R.D., Dasmahapatra, K.K., Kronforst, M., Joron, M., Jiggins, C.D., Mcmillan, W.O., Di Palma, F., Blumberg, A.J., Wakeley, J., Jaffe, D., Mallet, J., 2019. Genomic architecture and introgression shape a butterfly radiation. Science 366, 594–599.
Fritsch, P.W., 2001. Phylogeny and biogeography of the flowering plant genus *Styrax* (Styracaceae) based on chloroplast DNA restriction rites and DNA sequences of the internal transcribed spacer region. Mol. Phylogenet. Evol. 19, 387–408.
Frodin, D.G., Govaerts, R., 2003. World Checklist and Bibliography of Araliaceae: 1–444. Royal Botanic Gardens, Kew.
Gallego-Narbón, A., Wen, J., Liu, J., Valcárcel, V., 2022. Hybridization and genome duplication for early evolutionary success in the Asian Palmate group of Araliaceae. J. Syst. Evol. 60, 1303–1318.
Gatesy, J., Springer, M.S., 2014. Phylogenetic analysis at deep timescales: Unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80, 231–266.
Graham, A., 1993. History ofthe vegetation: Cretaceous (Maastrichtian)–Tertiary. In Flora of North America, Flora of North America Editorial Committee, Ed. Oxford Univ. Press, New York. 1, 57–70.
Grisebach, A., 1872. Die Vegetation der Erde nach Ihrer Klimatischen Anordnung. W. Engelmann, Leipzig.
Guo, Z.T., Sun, B., Zhang, Z.S., Peng, S.Z., Xiao, G.Q., Ge, J.Y., Hao, Q.Z., Qiao, Y.S., Liang, M.Y., Liu, J.F., Yin, Q.Z., and Wei, J.J., 2008. A major reorganization of Asian climate by the early Miocene. *Climate of the Past* 4:153-174.
Hardig, T.M., Soltis, P.S., Soltis, D.E., 2000. Diversification of the north american shrub genus *Ceanothus* (Rhamnaceae): Conflicting phylogenies from nuclear ribosomal DNA and chloroplast DNA. Am. J. Bot. 87, 108–123.
Harms, H., 1897. Zur Kenntnis der Gattungen *Aralia* und *Panax*. Botanische Jahrbucher für Systematik Pflanzeneschichte 23, 1–23.
Herrando-Moraira, S., Calleja, J.A., Galbany-Casals, M., Garcia-Jacas, N., Liu, J.Q., López-Alvarado, J., López-Pujol, J., Mandel, J.R., Massó, S., Montes-Moreno, N., Roquet, C., Sáez, L., Sennikov, A., Susanna, A., Vilatersana, R., 2019. Nuclear and plastid DNA phylogeny of tribe Cardueae (Compositae) with hyb-seq data: A new subtribal classification and a temporal diversification framework. Mol. Phylogenet. Evol. 137, 313–332.
Heyduk, K., Trapnell, D.W., Barrett, C.F., Leebens-Mack, J., 2015. Phylogenomic analyses of species relationships in the genus Sabal (Arecaceae) using targeted sequence capture. Biol. J. Linn. Soc. 117, 106–120.
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. https://doi.org/10.1016/j.ympev.2021.107118.
Hodel, R.G.J., Zimmer, E.A., Liu, B.B., Wen, J., 2022. Synthesis of nuclear and chloroplast data combined with network analyses supports the polyploid origin of the apple tribe and the hybrid origin of the Maleae-Gillenieae clade. Front. Plant Sci. 12.
Hoo, G., Tseng, C.J., 1978. Araliaceae. In: Flora Reipublicae Popularis Sinicae. Science Press, Beijing, p. 54.
Huang, T.C.m 1993. *Aralia* L.. In: Huang, TC ed. *Flora of Taiwan*. Editorial Committee of the Flora of Taiwan, Taipei. 3: 986-991.
Huelsenbeck, J.P., Rannala, B., Masly, J.P., 2000. Accommodating phylogenetic uncertainty in evolutionary studies. Science 288, 2349–2350.
Jansen, R.K., Ruhlman, T.A., 2012. Plastid genomes of seed plants, Genomics of chloroplasts and mitochondria. Springer: 103-126.
Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J., Wickett, N.J., 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. Appl. Plant Sci. 4, 1600016.
Junier, T., Zdobnov, E.M., 2010. The newick utilities: High-throughput phylogenetic tree processing in the unix shell. Bioinformatics 26, 1669–1670.
Kim, N.H., Jayakodi, M., Lee, S.C., Choi, B.S., Jang, W., Lee, J., Kim, H.H., Waminal, N. E., Lakshmanan, M., van Nguyen, B., Lee, Y.S., Park, H.S., Koo, H.J., Park, J.Y., Perumal, S., Joh, H.J., Lee, H., Kim, J., Kim, I.S., Kim, K., Koduru, L., Kang, K.B., Sung, S.H., Yu, Y., Park, D.S., Choi, D., Seo, E., Kim, S., Kim, Y.C., Hyun, D.Y., Park, Y.I., Kim, C., Lee, T.H., Kim, H.U., Soh, M.S., Lee, Y., In, J.G., Kim, H.S., Kim, Y.-M., Yang, D.C., Wing, R.A., Lee, D.Y., Paterson, A.H., Yang, T.J., 2018. Genome and evolution of the shade-requiring medicinal herb *Panax ginseng*. Plant Biotechnol. J. 16, 1904–1917.
Kitagawa, M., 1935. Contribution ad cognitionem florae Manshuricae. VI. *The Botanical Magazine*. 49, 222–233.
Kleinkopf, J.A., Roberts, W.R., Wagner, W.L., Roalson, E.H., 2019. Diversification of Hawaiian *Cyrtandra* (Gesneriaceae) under the influence of incomplete lineage sorting and hybridization. J. Syst. Evol. 57, 561–578.
Lanfear, R., Senfeld, T., Frandsen, P.B., Wright, A.M., Calcott, B., 2016. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357.
Lee, J.S., Kim, S.-H., Lee, S., Maki, M., Otsuka, K., Kozhevnikov, A.E., Kozhevnikova, Z. V., Wen, J., Kim, S.-C., 2019. New insights into the phylogeny and biogeography of subfamily Orontioideae (Araceae). J. Syst. Evol. 57, 616–632.
Leebens-Mack, J.H., Barker, M.S., Carpenter, E.J., Deyholos, M.K., Gitzendanner, M.A., Graham, S.W., Grosse, I., Li, Z., Melkonian, M., Mirarab, S., Porsch, M., Quint, M., Rensing, S.A., Soltis, D.E., Soltis, P.S., Stevenson, D.W., Ullrich, K.K., Wickett, N.J., DeGironimo, L., Edger, P.P., Jordon-Thaden, I.E., Joya, S., Liu, T., Melkonian, B., Miles, N.W., Pokorny, L., Quigley, C., Thomas, P., Villarreal, J.C., Augustin, M.M., Barrett, M.D., Baucom, R.S., Beerling, D.J., Benstein, R.M., Biffin, E., Brockington, S. F., Burge, D.O., Burris, J.N., Burris, K.P., Burtet-Sarramegna, V., Caicedo, A.L., Cannon, S.B., Çebi, Z., Chang, Y., Chater, C., Cheeseman, J.M., Chen, T., Clarke, N. D., Clayton, H., Covshoff, S., Crandall-Stotler, B.J., Cross, H., dePamphilis, C.W., Der, J.P., Determann, R., Dickson, R.C., Di Stilio, V.S., Ellis, S., Fast, E., Feja, N., Field, K.J., Filatov, D.A., Finnegan, P.M., Floyd, S.K., Fogliani, B., García, N., Gâteblé, G., Godden, G.T., Goh, F., Greiner, S., Harkess, A., Heaney, J.M., Helliwell, K.E., Heyduk, K., Hibberd, J.M., Hodel, R.G.J., Hollingsworth, P.M., Johnson, M.T.J., Jost, R., Joyce, B., Kapralov, M.V., Kazamia, E., Kellogg, E.A., Koch, M.A., Von Konrat, M., Könyves, K., Kutchan, T.M., Lam, V., Larsson, A., Leitch, A.R., Lentz, R., Li, F.-W., Lowe, A.J., Ludwig, M., Manos, P.S., Mavrodiev, E., McCormick, M.K., McKain, M., McLellan, T., McNeal, J.R., Miller, R.E., Nelson, M.N., Peng, Y., Ralph, P., Real, D., Riggins, C.W., Ruhsam, M., Sage, R.F., Sakai, A.K., Scascitella, M., Schilling, E.E., Schlösser, E.-M., Sederoff, H., Servick, S., Sessa, E.B., Shaw, A.J., Shaw, S.W., Sigel, E.M., Skema, C., Smith, A.G., Smithson, A., Stewart, C. N., Stinchcombe, J.R., Szövényi, P., Tate, J.A., Tiebel, H., Trapnell, D., Villegente, M., Wang, C.-N., Weller, S.G., Wenzel, M., Weststrand, S., Westwood, J. H., Whigham, D.F., Wu, S., Wulff, A.S., Yang, Y., Zhu, D., Zhuang, C., Zuidof, J., Chase, M.W., Pires, J.C., Rothfels, C.J., Yu, J., Chen, C., Chen, L., Cheng, S., Li, J., Li, R., Li, X., Lu, H., Ou, Y., Sun, X., Tan, X., Tang, J., Tian, Z., Wang, F., Wang, J., Wei, X., Xu, X., Yan, Z., Yang, F., Zhong, X., Zhou, F., Zhu, Y., Zhang, Y., Ayyampalayam, S., Barkman, T.J., Nguyen, N.-p., Matasci, N., Nelson, D.R., Sayyari, E., Wafula, E.K., Walls, R.L., Warnow, T., An, H., Arrigo, N., Baniaga, A.E., Galuska, S., Jorgensen, S.A., Kidder, T.I., Kong, H., Lu-Irving, P., Marx, H.E., Qi, X., Reardon, C.R., Sutherland, B.L., Tiley, G.P., Welles, S.R., Yu, R., Zhan, S., Gramzow, L., Theißen, G., Wong, G.-K.-S., Thousand, O., Plant, T.I., 2019. One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574, 679–685.
Lee-Yaw, J.A., Grassa, C.J., Joly, S., Andrew, R.L., Rieseberg, L.H., 2019. An evaluation of alternative explanations for widespread cytonuclear discordance in annual sunflowers (*Helianthus*). New Phytol. 221, 515–526.
Li, H.L., 1952. Floristic relationships between Eastern Asia and eastern North America. Trans. Am. Philos. Soc. 42, 371429.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with burrows–wheeler transform. Bioinformatics 25, 1754–1760.
Li, Z.Z., Gichira, A.W., Muchuku, J.K., Li, W., Wang, G.X., Chen, J.M., 2021. Plastid phylogenomics and biogeography of the genus *Monochoria* (Pontederiaceae). J. Syst. Evol. 59, 1027–1039.
Li, J., Stukel, M., Bussies, P., Skinner, K., Lemmon, A.R., Lemmon, E.M., Brown, K., Bekmetjev, A., Swenson, N.G., 2019. Maple phylogeny and biogeography inferred from phylogenomic data. J. Syst. Evol. 57, 594–606.
Li, R., Wen, J., 2013. Phylogeny and biogeography of *Dendropanax* (Araliaceae), an Amphi-Pacific disjunct genus between tropical/subtropical Asia and the Neotropics. Syst. Bot. 38, 536–551.
Li, H.L., 1972. Eastern Asia - Eastern North America speciespairs in wideranging genera. *In Floristics and Paleofloristics of Asia and Eastern North America*, ed. A. Graham, 65-78.
Liu, B.B., Hong, D.Y., Zhou, S.L., Xu, C., Dong, W.P., Johnson, G., Wen, J., 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, J.Q., Li, J.L., Lai, Y.J., 2021b. Plant diversity and ecology on the Qinghai-Tibet Plateau. J. Syst. Evol. 59, 1139–1141.
Liu, B.B., Ma, Z.Y., Ren, C., Hodel, R.G.J., Sun, M., Liu, X.Q., Liu, G.N., Hong, D.Y., Zimmer, E.A., Wen, J., 2021a. 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., Hodel, R.G.J., Xu, C., He, J., Zhou, W.B., Huang, C.H., Ma, H., Qian, G.Z., Hong, D.Y., Wen, J., 2022. 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. 00.
Lowry II, P.P., Plunkett, G.M., 2020. Resurrection of the genus *Heptapleurum* for the Asian clade of species previously included in *Schefflera* (Araliaceae). Novon, J. Bot. Nomen. 28, 143–170.
Lowry, P.P., Plunkett, G.M., Wen, J., Van Wyk, B.E., Tilney, P.M., 2004. Generic relationships in Araliaceae: Looking into the crystal ball. S. Afr. J. Bot. 70, 382–392.
Ma, Z.Y., Wen, J., Ickert-Bond, S., Nie, Z.L., Chen, L.Q., Liu, X.Q., 2018a. Phylogenomics, biogeography, and adaptive radiation of grapes. Mol. Phylogenet. Evol. 129, 258–267.

Ma, Z.Y., Wen, J., Tian, J.P., Jamal, A., Chen, L.Q., Liu, X.Q., 2018b. Testing reticulate evolution of four Vitis species from East Asia using restriction-site associated DNA sequencing. J. Syst. Evol. 56, 331–339.
Ma, ZY, Nie, ZL, Liu, XQ, Tian, JP, Zhou, YF, Zimmer, EA, Wen, J, 2023. Phylogenetic relationships, hybridization events, and drivers of diversification of East Asian wild grapes as revealed by phylogenomic analyses. Journal of Systematics and Evolution. https://doi.org/10.1111/jse.12918.
Ma, Z.Y., Nie, Z.L., Ren, C., Liu, X.Q., Zimmer, E.A., Wen, J., 2021. Phylogenomic relationships and character evolution of the grape family (Vitaceae). Mol. Phylogenet. Evol. 154, 106948.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.
Mai, U., Mirarab, S., 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. BMC Genomics 19 (S5), 272.
Mao, K.S., Wang, Y., Liu, J.Q., 2021. Evolutionary origin of species diversity on the Qinghai-Tibet Plateau. J. Syst. Evol. 59, 1142–1158.
Medina, R., Johnson, M.G., Liu, Y., Wickett, N.J., Shaw, A.J., Goffinet, B., 2019. Phylogenomic delineation of *Physcomitrium* (Bryophyta: Funariaceae) based on targeted sequencing of nuclear exons and their flanking regions rejects the retention of *Physcomitrella*, *Physcomitidrium*, and *Aphanorrhegma*. J. Syst. Evol. 57, 404–417.
Mirarab, S., Warnow, T., 2015. Astral-II: Coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31, i44–i52.
Mitchell, A., Li, R., Brown, J.W., Schönberger, I., Wen, J., 2012. Ancient divergence and biogeography of *Raukaua* (Araliaceae) and close relatives in the southern hemisphere. Aust. Syst. Bot. 25, 432–446.
Moore, A.J., Vos, J.M.D., Hancock, L.P., Goolsby, E., Edwards, E.J., 2017. Targeted enrichment of large gene families for phylogenetic inference: Phylogeny and molecular evolution of photosynthesis genes in the *Portullugo* clade (Caryophyllales). Syst. Biol. 67, 367–383.
Morales-Briones, D.F., Liston, A., Tank, D.C., 2018. Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the neotropical genus *Lachemilla* (Rosaceae). New Phytol. 218, 1668–1684.
Morales-Briones, D.F., Kadereit, G., Tefarikis, D.T., Moore, M.J., Smith, S.A., Brockington, S.F., Timoneda, A., Yim, W.C., Cushman, J.C., Yang, Y., 2021. Disentangling Sources of Gene Tree Discordance in Phylogenomic Data Sets: Testing Ancient Hybridizations in Amaranthaceae s.l. Syst. Biol. 70, 219–235.
Nakamura, T., Yamada, K.D., Tomii, K., Katoh, K., 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. Bioinformatics 34, 2490–2492.
Nussbaumer, T., Martis, M.M., Roessner, S.K., Pfeifer, M., Bader, K.C., Sharma, S., Gundlach, H., Spannagl, M., 2013. MIPS PlantsDB: a database framework for comparative plant genome research. Nucleic Acids Res. 41, D1144–D1151.
Otofuji, Y., Matsuda, T., Nohda, S., 1985. Opening mode of the Japan Sea inferred from the palaeomagnetism of the Japan Arc. Nature 317, 603–604.
Panero, J.L., Freire, S.E., Espinar, L.A., Crozier, B.S., Barboza, G.E., Cantero, J.J., 2014. Resolution of deep nodes yields an improved backbone phylogeny and a new basal lineage to study early evolution of Asteraceae. Mol. Phylogenet. Evol. 80, 43–53.
Payseur, B.A., Rieseberg, L.H., 2016. A genomic perspective on hybridization and speciation. Mol. Ecol. 25, 2337–2360.
Pease, J.B., Brown, J.W., Walker, J.F., Hinchliff, C.E., Smith, S.A., 2018. Quartet sampling distinguishes lack of support from conflicting support in the green plant tree of life. Am. J. Bot. 105, 385–403.
Plunkett GM, Wen J, Lowry PP. 2004. Infrafamilial classifications and characters in araliaceae: Insights from the phylogenetic analysis of nuclear (its) and plastid (trnl-trnf) sequence data. *Plant Systematics and Evolution* 245:1-39.
Plunkett, G.M., Wen, J., Lowry, P.P., Mitchell, A.D., Henwood, M.J., Fiaschi, P., 2018. Araliaceae. In: Kadereit, J., Bittrich, V. (Eds.), Flowering Plants. Eudicots, The Families and Genera of Vascular Plants, Springer, Cham, p. 15.
Qi, X.S., Chen, C., Hans, P.C., Sakaguchi, S., Liu, Y.H., Tanaka, N., Sakio, H., Qiu, Y.X., 2012. Molecular data and ecological niche modelling reveal ahighly dynamic evolutionary history of the East Asian Tertiary relict *Cercidiphyllum* (Cercidiphyllaceae). *New Phytologist* 196: 617-630.
Qi, X.S., Na, Y., Comes, H.P., Sakaguchi, S., Qiu, Y.X., 2014. A strong 'filter'effect of the east China Sea land bridge for East Asia's temperateplant species: Inferences from molecular phylogeography andecological niche modelling of *Platycrater arguta* (Hydrangea-ceae). BMC Evol. Biol. 14, 41.
Qiu, Y.X., Fu, C.X., Comes, H.P., 2011. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol. Phylogenet. Evol. 59, 225–244.
Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in bayesian phylogenetics using tracer 1.7. Syst. Biol. 67, 901–904.
Rieseberg, L., Soltis, D., 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evolutionary Trends in Plants 5, 65–84.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Rose, J.P., Toledo, C.A.P., Lemmon, E.M., Lemmon, A.R., Sytsma, K.J., 2021. Out of sight, out of mind: widespread nuclear and plastid-nuclear discordance in the flowering plant genus *Polemonium* (Polemoniaceae) suggests widespread historical gene flow despite limited nuclear signal. Syst. Biol. 70, 162–180.
Salichos, L., Rokas, A., 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497, 327–331.
Salichos, L., Stamatakis, A., Rokas, A., 2014. Novel information theory-based measures for quantifying incongruence among phylogenetic trees. Mol. Biol. Evol. 31, 1261–1271.
Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.
Seehausen, O., 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19, 198–207.
Smith, A.C., 1944. Araliaceae. In: North American flora, 28B. New York Bot. Gard, Bronx, pp. 3–41.
Smith, S.A., O'Meara, B.C., 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28 (20), 2689–2690.
Smith, S.A., Moore, M.J., Brown, J.W., Yang, Y., 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15, 150.
Solís-Lemus, C., Ané, C., 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. PLoS Genet. 12, e1005896.
Solís-Lemus, C., Bastide, P., Ané, C., 2017. PhyloNetworks: A Package for Phylogenetic Networks. Mol. Biol. Evol. 34, 3292–3298.
Soltis, D.E., Kuzoff, R.K., 1995. Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). Evolution 49, 727–742.
Soltis, P.S., Soltis, D.E., 2009. The role of hybridization in plant speciation. Annu. Rev. Plant Biol. 60, 561–588.
Stamatakis, A., 2014. Raxml version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stephens, J.D., Rogers, W.L., Heyduk, K., Cruse-Sanders, J.M., Determann, R.O., Glenn, T.C., Malmberg, R.L., 2015. Resolving phylogenetic relationships of the recently radiated carnivorous plant genus *Sarracenia* using target enrichment. Mol. Phylogenet. Evol. 85, 76–87.
Sun, M., Soltis, D.E., Soltis, P.S., Zhu, X., Burleigh, J.G., Chen, Z., 2015. Deep phylogenetic incongruence in the angiosperm clade rosidae. Mol. Phylogenet. Evol. 83, 156–166.
Szafer, W., 1961. Miocene flora from Stare Gliwice in upper Silesia. Institut Geologiczny Prace 33, 1–205.
Valcárcel, V., Wen, J., 2019. Chloroplast phylogenomic data support eocene amphi-pacific early radiation for the asian palmate core Araliaceae. J. Syst. Evol. 57, 547–560.
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.
Villaverde, T., Pokorny, L., Olsson, S., Rincon-Barrado, M., Johnson, M.G., Gardner, E. M., Wickett, N.J., Molero, J., Riina, R., Sanmartin, I., 2018. Bridging the micro- and macroevolutionary levels in phylogenomics: Hyb-seq solves relationships from populations to species and above. New Phytol. 220, 636–650.
Wang, H.X., Moore, M.J., Barrett, R.L., Landrein, S., Sakaguchi, S., Maki, M., Wen, J., Wang, H.F., 2020. Plastome phylogenomic insights into the Sino-Japanese biogeography of *Diabelia* (Caprifoliaceae). J. Syst. Evol. 58, 972–987.
Wang, H.X., Morales-Briones, D.F., Moore, M.J., Wen, J., Wang, H.F., 2021. A phylogenomic perspective on gene tree conflict and character evolution in Caprifoliaceae using target enrichment data, with Zabelioideae recognized as a new subfamily. J. Syst. Evol. 59, 897–914.
Wang, J.J., Yang, Y.P., Sun, H., Wen, J., Deng, T., Nie, Z.L., Meng, Y., 2016. The biogeographic south-north divide of *Polygonatum* (Asparagaceae Tribe Polygonateae) within eastern Asia and its recent dispersals in the Northern Hemisphere. PLoS One 11 (11), e0166134.
Weitemier, K., Straub, S.C., Cronn, R.C., Fishbein, M., Schmickl, R., Mcdonnell, A., Liston, A., 2014. Hyb-seq: Combining target enrichment and genome skimming for plant phylogenomics. Appl. Plant Sci. 2, 1400042.
Welsh SL, Atwood ND. 1975. Araliaceae. In: Welsh SL, Atwood N D, Reveal J L. Endangered, threatened, extinct, endemic, and rare or restricted Utah vascular plants. *The Great Basin Naturalist*. 35: 327-376.
Wen, J., 1993. Generic delimitation of *Aralia* (Araliaceae). Brittonia 45 (1), 47–55.
Wen, J., 1999. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annu. Rev. Ecol. Syst. 30, 421–455.
Wen, J., 2011. Systematics and biogeography of *Aralia* L. (Araliaceae): Revision of *Aralia* sects. *Aralia*, *Humiles*, *Nanae*, and *Sciadodendron*. Contributions from the United States National Herbarium 57, 1–172.
Wen, J., Nie, Z., Ickert-Bond, S.M., 2016. Intercontinental disjunctions between eastern Asia and western North America in vascular plants highlight the biogeographic importance of the Bering land bridge from late Cretaceous to Neogene. *J. System. Evol.,* 54: 469-490.
Wen, J., Liu, J.Q., Ge, S., Xiang, Q.Y., Zimmer, E.A., 2015. Phylogenomic approaches to deciphering the tree of life. J. Syst. Evol. 53, 369–370.
Wen, J., Harris, A.J., Kalburgi, Y., Zhang, N., Xu, Y., Zheng, W., Ickert-Bond, S.M., Johnson, G., Zimmer, E.A., 2018. Chloroplast phylogenomics of the New World grape species (*Vitis*, Vitaceae). J. Syst. Evol. 56, 297–308.
Wen, J., Nowicke, J.W., 1999. Pollen ultrastructure of Panax (the ginseng genus, Araliaceae), an eastern asian and eastern northamerican disjunct genus. Am. J. Bot. 86, 1624–1636.
Wen, J., Shi, S., Jansen, R., Zimmer, E., 1998. Phylogeny and biogeography of *Aralia* sect. *Aralia* (Araliaceae). Am. J. Bot. 85, 866–875.
Wen, J., Plunkett, G.M., Mitchell, A., Wagstaff, S., 2001. The evolution of Araliaceae: A phylogenetic analysis based on its sequences of nuclear ribosomal DNA. Syst. Bot. 26, 144–167.
Wen, J., Xiong, Z., Nie, Z.L., Mao, L., Zhu, Y., Kan, X.Z., Ickert-Bond, S.M., Gerrath, J., Zimmer, E.A., Fang, X.D., 2013. Transcriptome sequences resolve deep relationships of the grape family. PLoS One 8, e74394.
Wen, J., Zhang, J.-Q., Nie, Z.-L., Zhong, Y., Sun, H., 2014. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. Front. Genet. 5 https://doi.org/10.3389/fgene.2014.00004.

Wen, J., Zimmer, E.A., 1996. Phylogeny and biogeography of Panax L. (the ginseng genus, Araliaceae): Inferences from its sequences of nuclear ribosomal DNA. Mol. Phylogenet. Evol. 6, 167–177.

Widhelm, T.J., Grewe, F., Huang, J.P., Mercado-Díaz, J.A., Goffinet, B., Lücking, R., Moncada, B., Mason-Gamer, R., Lumbsch, H.T., 2019. Multiple historical processes obscure phylogenetic relationships in a taxonomically difficult group (Lobariaceae, Ascomycota). Sci. Rep. 9, 1–16.

Wolfe, K.H., Li, W.H., Sharp, P.M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. PNAS 84, 9054–9058.

Wu, C.Y., Wu, S.G., 1996. A proposal for a new plant kingdom (realm), the E. Asiatic Kingdom, its delimitation and characteristics. In: Zhang AO, Wu SG eds. *Floristic characteristics and diversity of East Asian plants, Proceedings of the First International Symposium on Floristic Characteristics and Diversity of East Asian Plants*. Beijing: China Higher Education Press; Berlin: Springer-Verlag. 3-42.

Xiang, Y., Huang, C.H., Hu, Y., Wen, J., Li, S., Yi, T., Chen, H., Xiang, J., Ma, H., 2017. Evolution of Rosaceae fruit types based on nuclear phylogeny in the context of geological times and genome duplication. Mol. Biol. Evol. 34, 262–281.

Xiang, Q.Y., Soltis, D.E., Soltis, P.S., 1998. The eastern Asian and eastern and western North American floristic disjunction: congruent phylogenetic patterns in seven diverse genera. Mol. Biol. Evol. 10, 178–190.

Xiang, Q.B., Lowry, P.P., 2007. Araliaceae. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. V13. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press.

Xie, G.W., 1997. On phytogeographical affinities of the forest floras between East China and Japan. Chin. Geogr. Sci. 7, 236–242.

Xing, Y., Ree, R.H., 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. PNAS 114 (17), E3444–E3451.

Yang, Y., Smith, S.A., 2014. Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: Improving accuracy and matrix occupancy for phylogenomics. Mol. Biol. Evol. 31, 3081–3092.

Zhang, D., Li, W., Xia, E.H., Zhang, Q.J., Liu, Y., Zhang, Y., Tong, Y., Zhao, Y., Niu, Y.C., Xu, J.H., Gao, L.Z., 2017. The medicinal herb *Panax notoginseng* genome provides insights into gisenoside biosynthesis and genome evolution. Mol. Plant 10, 899–902.

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. Astral-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf. 19, 153.

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. PLoSOne 10 (12), e0144701.

Zhang, N., Wen, J., Zimmer, E.A., 2016. Another look at the phylogenetic position of the grape order Vitales: Chloroplast phylogenomics with an expanded sampling of key lineages. Mol. Phylogenet. Evol. 101, 216–223.

Zhou, W., Xiang, Q.-Y.(J), Wen, J., 2020. Phylogenomics, biogeography, and evolution of morphology and ecological niche of the eastern Asian - eastern North American *Nyssa* (Nyssaceae). *J. System. Evol.,* 58, 571-603.

Zhou, W., Harris, A.J., Xiang, Q.Y.(J), 2022. Phylogenomics and biogeography of *Torreya* (Taxaceae)—Integrating data from three organelle genomes, morphology, and fossils and a practical method for reducing missing data from RAD-seq. J. Syst. Evol. 60, 1241–1262.

Zhu, Y.C., 1989. In Chang, ZY ed. Pl, Medic. Chinae Bor.-Orient, p. 59.

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.

Zuo, Y.J., Wen, J., Zhou, S.L., 2017. Intercontinental and intracontinental biogeography of the eastern Asian - eastern North American disjunct *Panax* (the ginseng genus, Araliaceae), emphasizing its diversification processes in eastern Asia. Mol. Phylogenet. Evol. 117, 60–74.