

Original Article

Multiple molecular data provide new insights into phylogeny and historical biogeography of East Asian *Artemisia* L. (Asteraceae)

Jingya Yu^{1,2}, Yun Han^{1,2}, Mingze Xia³, Hao Xu^{1,2}, Shuang Han^{1,2}, Xiaoping Li^{1,2}, Yu Niu^{1,2}, Shilong Chen¹, Faqi Zhang^{1,4,5,*} 

¹Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology & Institute of Sanjiangyuan National Park, Chinese Academy of Sciences, Xining 810008, China

²University of Chinese Academy of Sciences, Beijing 100039, China

³School of Pharmacy, Shandong Second Medical University, Weifang, 261053, China

⁴Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Xining 810008, China

⁵Xining Botanical Garden, Xining 810008, China

*Corresponding author. Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology & Institute of Sanjiangyuan National Park, Chinese Academy of Sciences, Xining 810008, China. E-mail: fqzhang@nwipb.cas.cn

ABSTRACT

Artemisia L. is one of the most diverse genera in the Asteraceae, widely used in agriculture and medicine, with a giant range of complicated taxa. The task of establishing the phylogeny difficulties owing to the highly similar morphological characters. East Asia, a biodiversity hotspot and major usage area for *Artemisia*, has received limited attention. Here, we collected 71 species (two subspecies, 94 samples) of *Artemisia* and its allies from East Asia and its neighbouring regions, and combined with public databases, ensuring representation of all East Asian subgenera. The phylogeny and historical biogeography of *Artemisia* and its allies in East Asia were reconstructed using plastome, nuclear ribosomal DNA (nrDNA), and nuclear single nucleotide polymorphism data obtained by genome skimming technology. Under the phylogenetic framework, we inferred introgression, divergence, and historical biogeography. We reveal strong nucleoplasmic conflicts in *Artemisia*, its allies, and subgenera. Past classifications could not classify most subgenera under *Artemisia* as monophyletic. East Asian *Artemisia* was probably diversified *in situ* in the Early Oligocene with the influences of climatic oscillations and geographic activities. *A. pectinate*, *A. palustris*, *A. keiskeana*, and *A. hedinii* appeared to belong to the early divergent lineages of modern East Asian *Artemisia*. These findings provide new insights into the evolution of *Artemisia*.

Keywords: *Artemisia*; phylogeny; plastome; ribosomal DNA; SNP

INTRODUCTION

Artemisia L., with >500 distinct or subspecific species, belongs to the Asteraceae family, which contains annuals, biennials, and perennials (Vallès and McArthur 2001, Vallès and Garnatje 2005). Most species in *Artemisia* are wind-pollinated (Wang 2004) and are found mainly in temperate arid and semi-arid regions of the Northern Hemisphere mid to high latitudes. Most members of this genus are of high economic value and are widely used in pharmaceuticals, fodder, pesticides, etc. (Obolskiy *et al.* 2011, Kapepula *et al.* 2020). Based on published fossil records and phylogeographic data, *Artemisia* was considered to be

originated in arid or subgrid habitats of temperate Asia during the Late Eocene or mid-Tertiary (Wang 2004, Wu *et al.* 2011). The modern distribution of *Artemisia* is characterized by centres of diversity in temperate and cold temperate regions of Eurasia, North America, and Asia (Lin 1993).

Because *Artemisia*'s remarkable diversification has resulted in a diverse and complicated range of species, the genus has been the subject of taxonomic and phylogenetic research since De Tournefort (1719) to the present day. A plethora of studies have delved into the multifaceted nature of *Artemisia*, encompassing its ecological niches, chemical compounds, cytological features,

pharmacological properties, and systematic relationships (Hayat et al. 2009, Abad et al. 2012). Despite this concern, there exist several unresolved challenges pertaining to the taxonomic classification of the genus. The delineation and elucidation of *Artemisia*'s phylogeny are fraught with complexity due to the high morphological diversity, the presence of intermediate forms, the occurrence of natural hybrids, and the prevalence of polyploidy (Riggins and Seigler 2012). These factors contribute to the intricate puzzle that is the evolutionary history of *Artemisia*, a genus that continues to defy easy categorization and understanding.

The monophyly of *Artemisia* has long been a contentious issue in botanical taxonomy. Multiple studies have provided evidence for a monophyletic lineage of the genus based on various factors such as sequence characters [internal transcribed spacer (ITS), external transcribed spacer region (ETS), nuclear genes, and some plastid genes] as well as pollen morphology (Torrell et al. 1999, Watson et al. 2002, Sanz et al. 2008, Malik et al. 2017, Hussain et al. 2019). However, recent molecular phylogenetic investigations have indicated the potential polyphyly of *Artemisia* (Pellicer et al. 2011, Sanz et al. 2011, Riggins and Seigler 2012, Yu et al. 2022, Jiao et al. 2023), based on ITS, ETS, plastome, and single nucleotide polymorphism (SNP). These investigations also documented potential homologies between *Artemisia* and its allies, and acknowledged the necessity for future investigation into the link between *Artemisia* and allies within the broader assemblage.

Hobbs and Baldwin (2013) proposed a classification for *Artemisia*, dividing it into six subgenera: *Absinthium*, *Artemisia*, *Dracunculus*, *Seriphidium*, *Tridantatae*, and *Pacifica*. However, it is worth noting that this classification may not accurately reflect the natural taxa within the genus. The majority of studies indicate that most subgenera were polyphyly, such as *Dracunculus*, *Absinthium*, *Seriphidium*, *Artemisia*, and *Tridantatae* (Garcia et al. 2011, Pellicer et al. 2011, Sanz et al. 2011, Hussain et al. 2019, Jiao et al. 2023). The latest molecular phylogenetic study has resulted in a new classification and expansion of *Artemisia* into eight subgenera (adding the subgenera *Pectinata* and *Ponticae*) based on SNP data from the one-to-one orthologues (Jiao et al. 2023). Despite this novel classification, further empirical data are imperative to substantiate its validity. There is also a notable disparity between the outcomes of morphological and molecular phylogenetic studies in determining the ancestral taxa of *Artemisia* (Hall and Clements 1923, Ling 1995, Riggins and Seigler 2012, Jiao et al. 2023). Hence, additional verification is required to establish the taxonomic and evolutionary lineage of *Artemisia* and its subgenera.

Plastomes contain important genetic information that can be used to resolve phylogenetic relationships in species with complex genomes. It has been successfully applied in the determination of phylogenetic relationships in complex taxa such as *Allium* L. (Xie et al. 2020), Leguminosae Juss. (Zhang et al. 2020b), subtribe Melocanninae of Poaceae Barnhart (Zhou et al. 2022), and *Rheum* L. (Zhang et al. 2022). The publication of a large number of plastomes (Shen et al. 2017, Kim et al. 2020, Shahzadi et al. 2020) and high-quality genomes (Liao et al. 2022, Miao et al. 2022) of *Artemisia* has provided a wealth of useful information on phylogeny. Nevertheless, there is still a scarcity of large-scale datasets with a rich phylogenetic signal to resolve

the phylogeny of *Artemisia* and to validate the taxonomic classifications.

Asia is one of *Artemisia*'s diversity hotspots, as well as a main distribution location for the genus' Old World species, containing ~300 species (Ling 1995). East Asia is one of the sub-hotspots for the distribution of *Artemisia*, with >200 species, nearly half of which are endemic (Ling 1995, Li 2007, Kim et al. 2020). Previous molecular phylogenetic investigations have encompassed a substantial quantity of *Artemisia* from East Asia, but their phylogenetic relationships remain controversial (Riggins and Seigler 2012, Hussain et al. 2019, Kim et al. 2020, Jiao et al. 2023). Meanwhile, biogeography and evolution of *Artemisia*, particularly in this region, remains less studied in depth.

East Asia comprises the 'high step', which includes the Qinghai-Tibetan Plateau (QTP), the Himalayas, and the Hengduan Mountains (Wang et al. 2012), as well as the 'low step', which includes eastern China, the Japanese archipelago, and the Korean Peninsula (Qiu et al. 2009). The current distribution pattern of the East Asian flora has been influenced by geologic history, floristic components, and changes in plant diversity and competition. This has resulted in varying species composition, endemism, and richness across different floristic regions (Wu et al. 2011, Chen et al. 2018a). Therefore, we collected *Artemisia* species from East Asia and its neighbouring regions, and combined with published data from the GenBank database, the final dataset includes all *Artemisia* subgenera in the East Asian region. We attempted to investigate (i) the monophyly of *Artemisia* in the East Asian region and the phylogenetic relationships of the genus and its close relatives; (ii) the applicability of current taxonomic divisions under *Artemisia*; (iii) phylogenetic conflicts between different datasets (plastome, nuclear ribosomal DNA (nrDNA), and genome-wide SNPs); and (iv) the origin and historical biogeography of the genus *Artemisia* in the East Asian region. Our study provides a reference for understanding *Artemisia* phylogeny and guiding the selection of datasets in future phylogenetic studies.

MATERIALS AND METHODS

Sample collection and DNA extraction

We collected fresh leaves from a total of 94 individuals in the field (Supporting Information, Table S1), including 64 species of *Artemisia*, six species of *Ajania* and one species of *Brachanthemum*. To show the relationships within the species, we included two or more individuals from different localities for some species. Voucher specimens for the species sampled were deposited in the QTP Museum of Biology (HNWP), Northwest Institute of Plateau Biology, Chinese Academy of Sciences.

We extracted total genomic DNA using ~10 mg of silica-dried leaf tissue through modified CTAB protocols (Doyle 1987). The extracted DNAs of all the individuals were then sent to Novogene (Beijing, China). The genomic DNA library was generated using NEB Next[®] Ultra[™] DNA Library Prep Kit for Illumina (NEB, USA) following the manufacturer's recommendations, and index codes were added to each sample and sequenced on an Illumina HiSeq 2500 sequencer (San Diego, CA, USA) using the paired-end option (2 × 150 bp). To include more comprehensive information on East Asian *Artemisia*, we

downloaded 43 plastomes of *Artemisia* and its relative genus from GenBank (Supporting Information, Table S2, <https://www.ncbi.nlm.nih.gov/genbank/>). A total of 137 individuals were included in this study, of which 78 species and 119 individuals were from *Artemisia*, seven species and eight individuals were from *Ajania*, seven species and eight individuals were from *Chrysanthemum*, one was from *Neopallasia pectinata*, and one was from *Brachanthemum pulvinatum*. Here, *B. pulvinatum* served as an outgroup of *Artemisia* and its relative genus.

Assembly and phylogenetic datasets

Data from sequencing were quality and filtered using FastP v.0.23.2 (Chen *et al.* 2018b). The plastome and complete nrDNA sequences (contain 18S-ITS1-5.8S-ITS2-26S) of the newly sampled species were then *de novo* assembled using GetOrganelle v.1.7.5 (Jin *et al.* 2020).

These plastomes were annotated and examined using the published plastome of *A. tangutica* (GenBank accession MT701043) as a reference, which was assembled using Illumina and PacBio sequencing data. For both the assembled and GenBank-downloaded plastomes, we used PGA (Qu *et al.* 2019) and the online tool CPGAVAS2 (Shi *et al.* 2019) for annotation to ensure consistency and accuracy of the gene coding sequence (CDS) data in the analysis.

PhyloSuite v.1.2.2 (Zhang *et al.* 2020a) was used to extract protein coding genes (CDSs) sequences. As for the analysis of plastic phylogeny, were constructed three phylogenetic datasets: Dataset I with 80 shared CDSs; Dataset II with 80 CDSs first and second codons concatenated (CDS¹⁺²); and Dataset III containing the whole plastome.

There were 98 nrDNAs assembled, of which 94 were newly sequenced in our study and the other four were assembled from the published sequencing data. The Sequence Read Archive numbers were displayed in Supporting Information, Table S3.

The published nrDNA data of *Atractylodes chinensis* (GenBank accession MZ456956), a closely related species of *Artemisia*, was used as a reference sequence for manual annotation and examination of the newly obtained nrDNA from this study. The annotated nrDNA sequences were merged as Dataset IV. To correspond to species in Dataset IV, Dataset III was abridged and reconstituted as Dataset V.

Phylogenetic analysis

The multiple sequence alignment was performed using MAFFT v.7.505 (Katoh and Standley 2013) for all datasets of the locus. The alignment parameters were set to 'G-INS-i (accurate)'. Dataset I used the 'codon' mode, while the remaining datasets were in the 'normal' mode. Substitutional saturation was assessed for each dataset in DAMBE v.7.0.68 (Xia and Xie 2001) and measured using the substitution saturation index (I_{ss}). No substitution saturation was seen in the data (Supporting Information, Table S5), hence all datasets collected in this study were included for further analysis.

Maximum likelihood (ML) method and Bayesian inference (BI) were applied to all the datasets to reconstruct the phylogenetic relationships of *Artemisia* and its related taxa. For ML analyses, the best-fit model was inferred using ModelFinder (Kalyaanamoorthy *et al.* 2017) based on Bayesian information

criterion. ML trees were run in IQ-TREE v.2.0.3 (Nguyen *et al.* 2015) under substitution models shown in Table S3 (in the Supporting Information), and 1 000 000 bootstrap replicates were applied in each single run. For Bayesian inference, we used ModelFinder to infer the best-fit model based on the corrected Akaike information criterion (AICc) and subsequently performed BI tree inference using MrBayes v.3.2.7 (Ronquist *et al.* 2012). Each BI analysis was performed by two independent runs of four one million generations of Monte Carlo Markov chains (MCMCs) under the GTR+G+I substitution model, with trees sampled every 1000 generations. The first 25% of the sampled trees were burned and the rest were used to generate consensus trees and calculate Bayesian posterior probabilities. The trees were visualized and edited using Figtree and Interactive Tree of Life (iTOL) (Letunic and Bork 2019).

To evaluate the reliability of phylogenetic tree selection, we used CONSEL (Shimodaira and Hasegawa 2001) to analyse topologies that exhibit substantial conflicts. The outcomes derived from the distinct datasets were individually inputted into RAxML (Stamatakis 2014), and the likelihood values for each locus of the matrix were computed for the various topologies.

SNP calling and genetic structure analysis

A. annua chromosome-level genome (Liao *et al.* 2022) as the reference, Hisat2 (Kim *et al.* 2019) was used to align clean data to the reference genome with the parameter set to --no-spliced-alignment. The process of identifying genetic variants involved the use of picard v.2.27.4 (Broad Institute 2019) and gatk4 v.4.3.0.0 (McKenna *et al.* 2010) to perform SNP calling. The subcomponents of GATK4 VariantFiltration were used to exclude potential false-positive variant calls. The exclusion of these calls was carried out using the specified parameters: QD < 2.0 || FS > 60.0 || MQ < 40.0 || SOR > 3.0 || MQRankSum < -12.5 || ReadPosRandSum < -8.0. To enhance the manageability of the variant call format (VCF), we used vcftools v.0.1.16 (Danecek *et al.* 2011) to filter and format the variant data. The filtering criteria included SNP loci with a missing rate < 50%, a minor allele count > 3, and a minor allele frequency > 0.05. The home scripts were used to exclude variants that were found within a 7 bp range around indel locations and inside the intervals of repeat sequences. As a final filtering strategy, we used Plink v.1.90b6.21 (Purcell *et al.* 2007) and applied the parameters 'indep-pairwise 50 10 0.2' to obtain a subset of SNPs that were pruned for linkage disequilibrium. This subset was further filtered to include only SNPs with a minor allele frequency > 5%. The Dataset VI was composed of the final set of high-quality SNPs that were obtained.

The ML phylogenetic tree was constructed based on the genetic distance matrix among the total SNP data set from all 98 individuals using IQ-TREE and the tree was displayed using iTOL. For genetic structure analysis, ADMIXTURE v.1.3.0 (Alexander *et al.* 2009) was used to assess cross-validation error values with parameter settings from $K = 1$ to $K = 10$, and we chose $K = 7$ to $K = 9$ for visualization using the R package ggplot2.

Landscape tree analysis

To investigate the variability present in phylogenetic trees, the Robinson-Foulds algorithm (Robinson and Foulds 1981) were

used to analyse the statistical distribution of trees. Datasets I, II, and III consisted of 137 individuals, while Datasets IV, V, and VI contained 98 individuals. The distances between unrooted trees were computed using the R v.4.3.0 package TREESPACE v.1.0.0 (Jombart et al. 2017), following the workflow by Gonçalves et al. (2019). Subsequently, the first two principal coordinate analyses (PCoAs) were estimated. Results were visualized using ggplot2 v.3.4.3.

Test for reticulation

To investigate relationships and reticulation among East Asian *Artemisia*, we generated phylogenetic networks from the unlinked SNP matrix. Networks were constructed in SplitsTree4 v.4.17.1 (Huson and Bryant 2006) using default settings; implementing Uncorrected P/Hamming Distances to obtain distance matrices, NeighborNet to obtain splits.

The software tool Dsuite (Malinsky et al. 2021) facilitates the computation of the D and f_4 statistics using SNPs stored in VCF files. The program Fbranch has been incorporated into Dsuite to untangle interrelated f_4 -ratio outcomes and attribute indications of gene flow to particular branches within a phylogenetic tree. Here, we employed Dsuite to evaluate the extent of gene flow and introgression using the genomic dataset of VCF files containing biallelic genomic SNPs. The D and f_4 -ratio statistics (Patterson et al. 2012) were calculated for all trios using Dtrios. The f -branch statistics were conducted by using ML tree as input. The output f -branch statistics in matrix-like format were then displayed using the plotting function.

Divergence time estimation and ancestral range reconstruction

Previous studies have demonstrated that plastid-based phylogenies are more likely to reveal clear geographical patterns (Pham et al. 2017, Hodkinson et al. 2019, Yang et al. 2021). Therefore, the divergence times of *Artemisia* were estimated based on Dataset III with the Bayesian approach conducted in BEAST v.1.8.4 (Drummond and Rambaut 2007). The analysis employed the BEAST software, which used a relaxed molecular clock model and lognormally distributed substitution rates for each branch in the phylogenetic tree. According to the result of ModelFinder's calculations, the GTR+ Γ +I nucleotide substitution model was used, along with a birth–death incomplete sampling speciation process tree prior. The calibration of fossils was conducted by using evidence of the presence of *Artemisia* pollen fossils found in the Tethyan region (Wang 2004), which has been identified in Early Oligocene sediments. Based on fossil information, we used the lognormal prior, setting an offset of 31 Mya, a mean of 0.5, and a standard deviation of 0.5. The default values were assigned to all other preceding variables.

Two independent MCMC analyses were terminated at 30 million generations (one tree was preserved every 1000 generations) with a 10% ageing rate. Two independent runs were merged in LogCombiner v.1.6.2. Tracer v.1.7 (Rambaut et al. 2018) was used to check the convergence of the effective sample size. Maximum clade credibility tree with median heights was generated using TreeAnnotator v.1.10.4. FigTree (Rambaut 2007) was used to implement the visualization. To examine the possibility of rapid species differentiation in *Artemisia* during

its evolutionary history, we used the R package ape to generate lineage-through-time plots.

Following the studies of the flora of *Artemisia* and the flora of East Asia (Ling 1995, Wu et al. 2011, Chen et al. 2018a) *Artemisia* and its allies were divided into five regions: (A) China–Japan; (B) Hengduan–Himalaya; (C) QTP; (D) Tethyan; and (E) Pan-Arctic. Of these, (A)–(C) together constitute the East Asian flora. Species distribution information is referenced from the Global Biodiversity Information Facility (<https://www.gbif.org/>). Accounting for phylogenetic uncertainty and less comprehensive outgroups, BEAST-generated trees were removed species outside *Artemisia* for ancestral range reconstruction analysis. Modified trees were exported to RASP v.4.2 (Yu et al. 2015) and models were compared using the R package BioGeoBEARS (Matzke 2018) for the biogeographic model's DEC, DIVALIKE, and BAYAREALIKE with/without J-parameters. The optimal model was chosen by evaluating the Akaike information criterion values.

RESULTS

Plastome, ribosomal DNA, and SNP datasets

For the 94 newly sequenced individuals, we obtained a total of 1078.89 Gb of sequencing data, with an average of 11.47 Gb (Supporting Information, Table S1). The total sequence length of the plastome of *Artemisia* ranged from 149 849 bp (*A. eriopoda*) to 151 367 bp (*A. anomala*, Supporting Information, Table S3). *Artemisia*'s plastomes were conservative in structure, without obvious structural variation. The *Artemisia* plastomes all comprised a long single-copy sequence (ranging 81 609–83 083 bp), a short single-copy sequence (ranging 18 175–18 539 bp), and two inverted repeats (ranging 24 927–24 976 bp). The number of gene CDSs in the *Artemisia* plastomes was 132–133 (Supporting Information, Table S4). Matrix lengths of different datasets based on plastomes were 68 133 bp for Dataset I, 45 423 bp for Dataset II, 150 352 bp for Dataset III, and 150 335 bp for Dataset V.

We assembled a total of 98 nrDNAs, and the details are summarized in Supplementary Table S3 (in the Supporting Information). Lengths of the obtained nrDNA ranged from 4405 bp (*A. thellungiana*) to 5853 bp (*A. sieversiana*), and the length of the integration matrix for Dataset IV was 5836 bp. The suitability of these datasets for phylogeny showed that the I_{ss} values were much smaller than the I_{ss-c} values (Supporting Information, Table S5).

The paired-end reads obtained from a total of 98 individuals were mapped to the haploid reference genome of *A. annua*. Following the process of base calling, the overall dataset consisted of 545 919 806 sites. Subsequently, on applying filtering, the number of high-quality variant sites was determined to be 8604 for Dataset VI.

Plastid phylogeny

The plastid phylogeny comprised 137 individuals in all, of which 119 *Artemisia* individuals represent the five major subgenera (Hobbs and Baldwin 2013) or seven major subgenera (Jiao et al. 2023) in East Asia. Phylogenetic trees (Fig. 1 and Supporting Information, Figs S1–S5) constructed from the various datasets

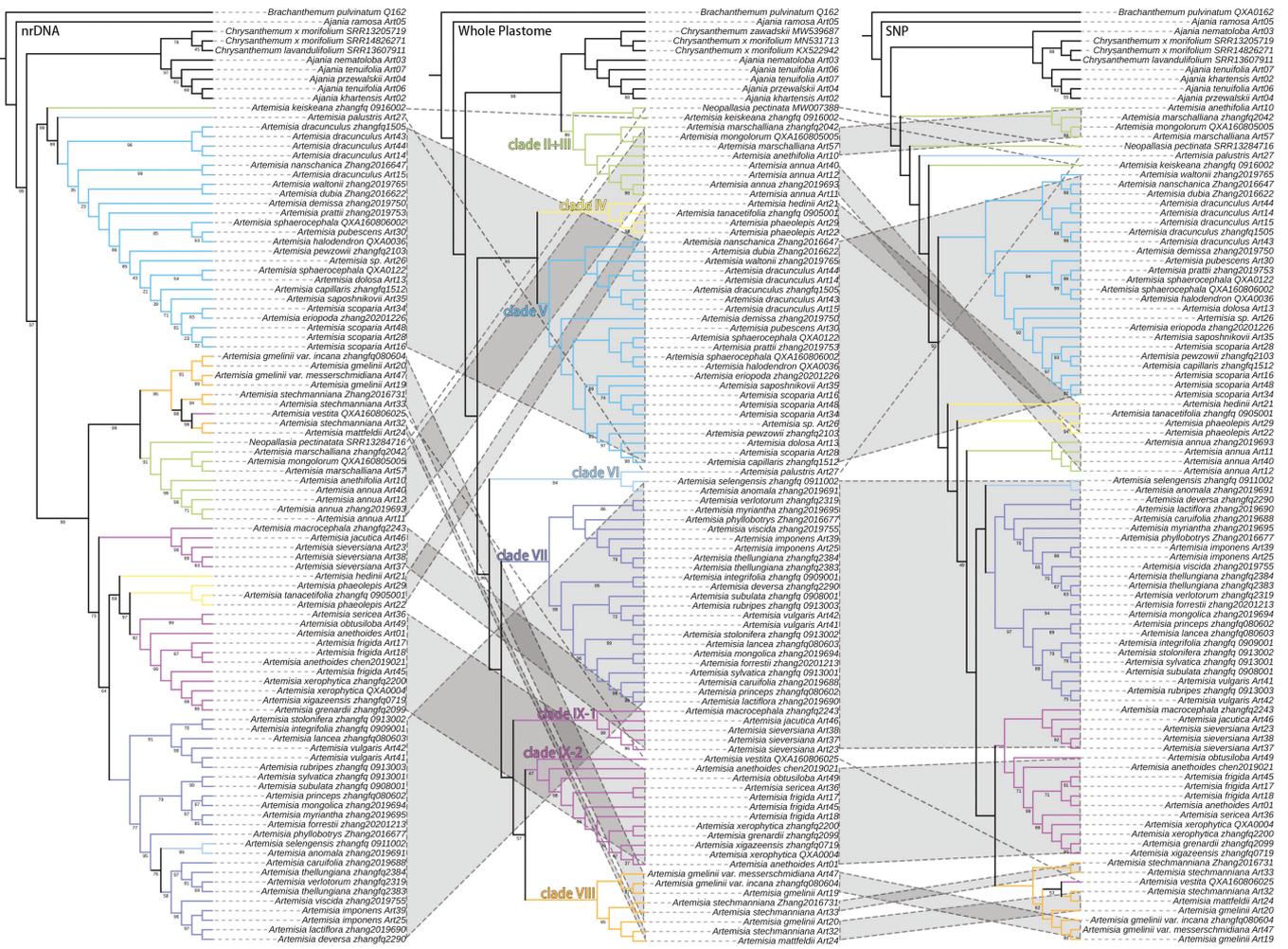


Figure 2. Comparison of nrDNA-based (left), plastome-wide (middle), and SNP (right) phylogenetic relationships. The values associated with branches are ML bootstrap values. Nodes with no numbers indicate 100% bootstrap support.

revealed almost similar topology. We showed the phylogeny built on the whole plastomes with better support (Fig. 1). Most branches of the phylogenetic trees had high support in our results. We compared the phylogenetic results with the main classification of *Artemisia* subgenera from previous studies (Ling 1995, Hobbs and Baldwin 2013, Jiao et al. 2023). However, neither infrageneric classification system was able to resolve the subgenera *Absinthium*, *Artemisia*, *Dracuncululus*, *Seriphidium*, *Pectinata*, and *Ponticea* into monophyletic taxa. We have divided *Artemisia* and its related taxa into nine lineages based on the plastid phylogeny. *Artemisia*, *Ajania*, and *Chrysanthemum* are strongly supported as sister genera. *Ajania* and *Chrysanthemum* have nested within Clade I. Clade II bootstrap BS = 100 and Clade III (bootstrap BS = 98) are complex, with species from multiple subgenera. According to Jiao et al.'s (2023) subgeneric classification of *Artemisia*, Clade V (bootstrap BS = 100) contains mainly species of the subg. *Dracuncululus*, Clade VII (bootstrap BS = 100) contains mainly species of the subg. *Artemisia*, clade VIII (bootstrap BS = 96) contains mainly species of the subg. *Ponticea*, and Clade IX (bootstrap BS = 84) contain mainly species of the subg. *Absinthium*.

We used PCoA to explore the inconsistency of the phylogenetic trees. The first and second axes of PCoA explained

42.3% and 21.8% of the variation in tree topology, respectively (Supporting Information, Fig. S8A). The CDS-based (Dataset I) topology is closer to those of the whole plastomes (Dataset III) than to those of CDS¹⁺². In the phylogenetic results constructed for the different datasets, the topologies of Clades II, IV, and VI conflict: (i) Clade II forms a sister with Clade I or Clade III; and (ii) Clade IV or Clade VI forms a sister with Clade V. We used CONSEL to validate conflicting clades and all datasets supported Clade II being a sister to Clade III and Clade IV being a sister to Clade V (Supporting Information, Table S6).

Comparison of phylogeny between plastome, NRDNA, and SNP

For 98 individuals containing raw sequencing data, we further compared phylogenetic relationships based on plastomes, nrDNA, and SNPs (Fig. 2 and Supporting Information, Figs S6, S7). There existed a notable incongruence between the topologies of the plastid and nuclear ML datasets. The three datasets recovered varying numbers of clades and demonstrated contradicting links among species and clades, which were statistically well-supported (Fig. 2). There was a strong nucleoplasmic conflict in *Artemisia*. The monophyly of *Artemisia* is supported by the phylogenetic analysis of nrDNA and SNP. However, plastid

phylogeny resolved *Artemisia* as a paraphyletic group. Except for Clades II + III and Clade VI, the members of the remaining clades has remained largely consistent, although they differ in position in the phylogeny based on different datasets.

For the phylogenetic trees based on 98 individuals, the first and second axes of PCoA explained 72.5% and 25.8% of the variation in tree topology, respectively (Fig. S8B in the Supporting Information). The topology based on nrDNA, whole plastid genomes, and SNPs is vastly different.

Genetics structure and reticulation

The genetic structure of the major lineages of *Artemisia* showed that some members of the major lineages exhibit varying degrees of admixture of genetic components, excepted for clade VIII, with a range of K of 7–9 (Supporting Information, Fig. S9). Clade III species at different locations in the SNP phylogeny showed almost uniform genetic components. *A. selengensis* and *A. anomala* in Clade VI showed a mixed type of Clades III and VII.

The NeighborNet of East Asian *Artemisia* showed extensive reticulation (Supporting Information, Fig. S10). There are clear cross links between Clades V, VII, and VIII, and these clades are taxonomically well differentiated. Notably, most lineages showed extensive reticulation before it diversified.

Low f_b values across numerous species suggested the occurrence of reticulate evolution on a broader temporal scale within the evolutionary framework. Multiple species pairs in *Artemisia* exhibited an abundance of shared alleles. However, it is noteworthy that significant introgression primarily took place within the major lineages, as depicted by the red dashed box in Figure 3. *N. pectinata*, *A. palustris*, *A. keiskeana*, and *A. hedinii* show more complex genetic components than most species of *Artemisia*, but we did not detect significant introgression in these species.

Divergence time estimation and ancestral range reconstruction

Molecular clock estimates indicated that the crown groups of *Artemisia* and its related taxa were estimated at 31.29 Mya [95% highest posterior density (HPD) 32.94–31.02 Mya] in the Early Oligocene (Fig. 4). The relatively early diverging clades under *Artemisia* are Clades II, III, IV, and VI, whose crown group times were inferred to be 22.25 Mya (HPD95% 26.47–16.82 Mya), 28.78 Mya (HPD95% 30.43–26.53 Mya), 21.74 Mya (HPD95% 26.01–16.44 Mya), and 25.18 Mya (HPD95% 27.95–21.90 Mya), respectively. Clade V underwent diversification at ~17.62 Mya (95%HPD 21.38–14.12 Mya). During the Early Miocene at 20.30 Mya (95%HPD 23.46–17.24 Mya), and Clade VII diverged from its sister clade and underwent diversification at 10.88 Mya (95%HPD 13.78–8.45 Mya). Clades VIII and IX diverged at 18.53 Mya (95% HPD 21.58–15.49 Mya). The lineage-through-time results indicated that *Artemisia* underwent a rapid lineage accumulation between the Pliocene and the Pleistocene (Supporting Information, Fig. S11).

We chose the BAYAREALIKE+J model for the ancestral range reconstruction (Supporting Information, Table S7). To explore the historical process of diversification of modern East Asian *Artemisia*, the ancestral distribution state of each of the major nodes was estimated based on their recent distributions

acquired from floristic works and specimen information (Fig. 3, Supporting Information Figs S12, S13). Modern *Artemisia* in East Asia was likely to diversified *in situ*, with a possible ancestral distribution in the China–Japan, followed by a spread to the Pan-Arctic and Tethyan regions. Notably, the two principal branches of Clade VII seem to have evolved in parallel, adapting to different altitudes.

DISCUSSION

Phylogenetic relationship of *Artemisia* and its related genera

Employing three distinct datasets, we have constructed a robust backbone phylogeny of *Artemisia* in East Asia, encompassing all subgenus under *Artemisia* in East Asia and the majority of lineages previously identified (Ling 1995, Hobbs and Baldwin 2013, Jiao *et al.* 2023). Our analysis reveals that, in contrast to the nrDNA phylogeny, the plastid and SNP phylogenies exhibit stronger support across the identified clades. The results showed high topological incoherence and severe nucleoplasmic conflicts between them. In addition, reticulate evolution such as hybridization, incomplete lineage, plastid capture events, or recombination may have emerged across *Artemisia* and its relatives.

Achieving congruent phylogenetic signals between cytoplasmic and nuclear genomes is often challenging, a phenomenon well-documented in the literature (Soltis *et al.* 1991, Liu *et al.* 2020). Hybridization, incomplete genealogical sorting (ILS), polyploidy, plastid capture, etc. may participate in nucleoplasmic conflicts (Degnan and Rosenberg 2009, Kawabe *et al.* 2018, Wang *et al.* 2018). ILS allows for the sharing of identical haplotypes across multiple lineages without the need for hybridization, leading to widespread genetic admixture across a taxon's range (Wu and Campbell 2005, Pasquet *et al.* 2021). However, Plastid capture commonly happens across or within species with sympatric distribution and reproductive compatibility (Liu *et al.* 2017), probably explaining the phylogenetic inconsistency between plastids and nuclear in this study. In this study, not only interspecies, but also intersubgenus and even intergeneric phylogenetic relationships showed strong nucleoplasmic conflicts (Fig. 2), suggesting that plastid capture events may be common in *Artemisia*. Reports of reproductive compatibility between *Artemisia* and *Chrysanthemum* (Zhu *et al.* 2013), along with overlapping distributions of *Artemisia* with *Ajania* in the QTP and Himalaya, and with *Chrysanthemum* in the China–Japan region, imply that hybridization events may have occurred in these areas, with introgression probably occurring early in their divergence history (Fig. 4, Supporting Information Figs S8, S12, S13). Plastid exchange between major lineages early in evolution may obscure phylogenetic relationships even at higher levels (Stull *et al.* 2020). This imply that non-monophyletic plastid phylogenies may be explained by plastid capture through hybridization during the early diversification of *Artemisia*. In addition, we support the taxonomic revision (Muldashev 1981, Huang *et al.* 2017) of *Ajania* and *Phaeostigma* (*Aj. ramosa* = *P. ramosa*, and *Aj. variifolium* = *P. variifolium*). Recent reticulation conflicts between *Ajania* and *Chrysanthemum* may have led to a blurring of the boundaries between the two genera (Supporting Information, Fig. S10).

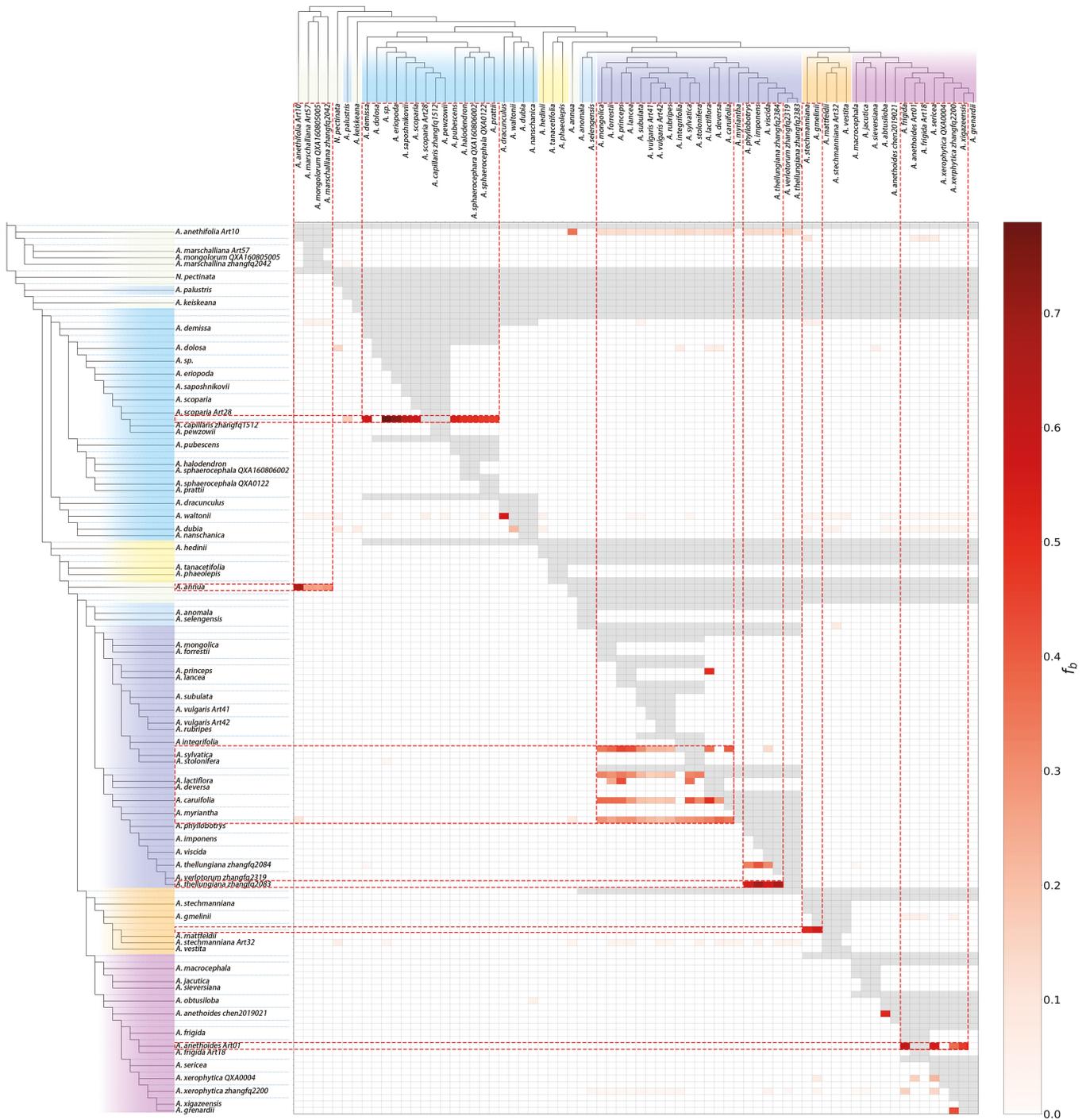


Figure 3. Results of Dsuite statistics. f -branch statistics showing deeper introgressions among taxa and lineages. Grey shadings indicate untestable pairs, given our sampling and the phylogenetic tree. Boxes indicated signals of introgression detected among clades.

Notably, recent molecular phylogenetic studies have identified *Artemisia* as non-monophyletic, with *A. stracheyi* and *Aj. quercifolia* clustered together as a single clade (Jiao et al. 2023). While the nuclear data from this study suggest that East Asian *Artemisia* is monophyletic, the absence of these key species means that a definitive conclusion on the monophyly of the genus cannot be drawn. Future studies will require more comprehensive sampling and accurate data for further study.

Phylogeny of the major lineages and species of East Asian *Artemisia*

Previous studies have made revisions to the classification of *Artemisia* (Hobbs and Baldwin 2013, Jiao et al. 2023). However, most of the subgenera lacked support as monophyletic groups. This study confirmed that the subg. *Dracunculus* is an early lineage of *Artemisia* (Sanz et al. 2008, 2011, Jiao et al. 2023), rather than the subg. *Absinthium* (Tkach et al. 2008a). The subg. *Dracunculus* (Clade V), subg. *Artemisia* (Clade VII), subg.

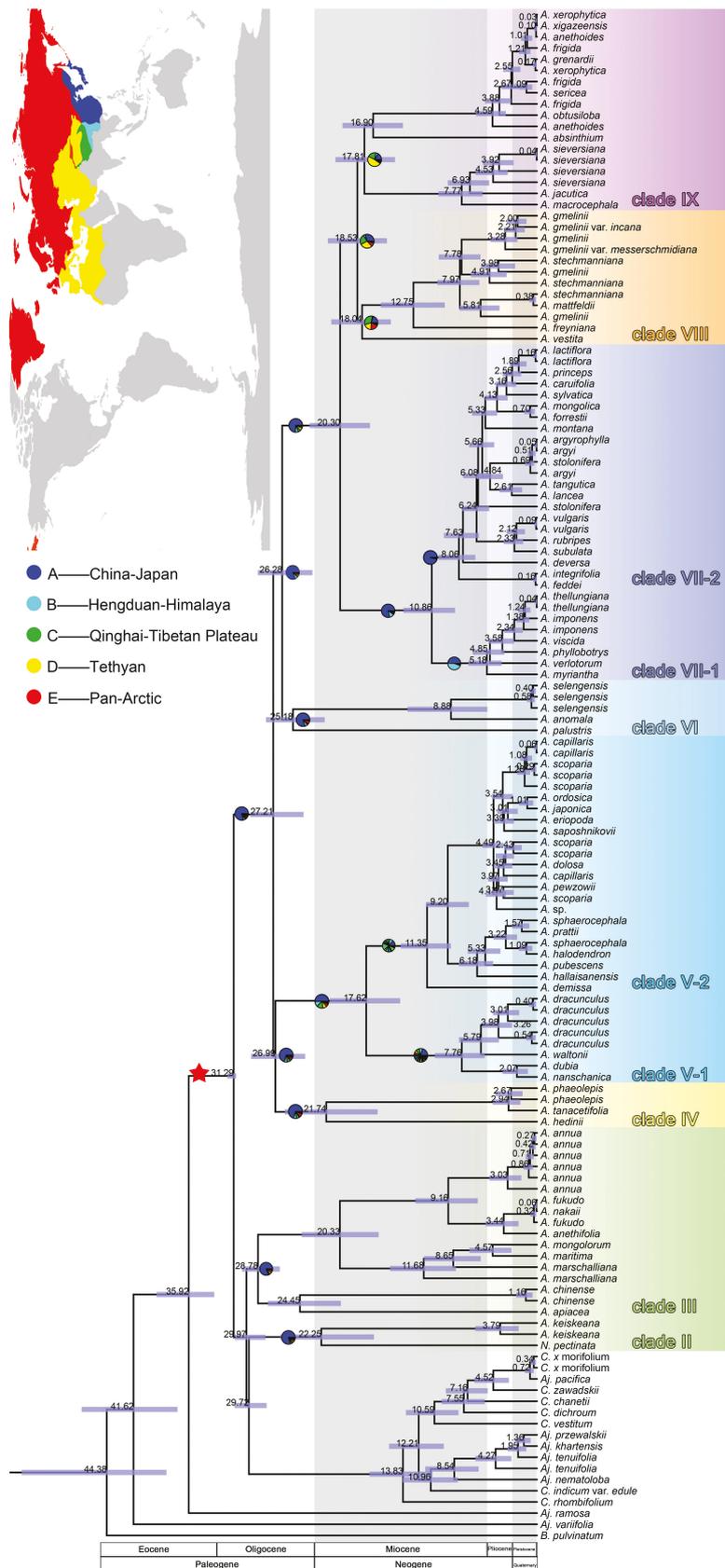


Figure 4. Divergence time estimates for *Artemisia* based on the whole plastomes. Numbers associated with branches represent the mean estimated divergence time (Mya) and the blue bars correspond to the 95% HPD of divergence time. The red star indicates one calibrating point. Pie charts show relative probabilities of estimated ancestral ranges from the BAYAREALIKE + J model: blue represents the China–Japan region; light blue represents the Hengduan–Himalaya region; green represents the QTP region; yellow represents the Tethyan region; and red represents the Pan-Arctic region.

Poniticae (Clade VIII), and subg. *Absinthium* (Clade IX) was more clearly divergent and was supported by genome-wide SNPs and plastid datasets (Fig. 2, Supporting Information Fig. S10). However, the phylogenetic positions of subg. *Artemisia*, subg. *Poniticae* and subg. *Absinthium* varied across datasets and studies, and can be categorized mainly into the following: (i) subg. *Artemisia* is the related taxa to subg. *Poniticae*, subg. *Absinthium*, and subg. *Seriphidium* (Malik et al. 2017, Jin et al. 2023); (ii) subg. *Artemisia* and subg. *Absinthium* are the sister to subg. *Poniticae* and subg. *Seriphidium* (Watson et al. 2002, Jiao et al. 2023); (iii) subg. *Absinthium* is the related taxa to subg. *Artemisia* and subg. *Seriphidium* (Sanz et al. 2011); and (iv) subg. *Poniticae* is the related taxa to subg. *Seriphidium* and subg. *Artemisia* (Sanz et al. 2008, Hobbs and Baldwin 2013). Among them, the phylogenetic results of this study based on the plastid and SNPs dataset are more supportive of (i), whereas the nrDNA datasets are more supportive of (ii). Subg. *Poniticae* and subg. *Absinthium* may share a common ancestor (Figs 1, 2, Supporting Information Fig. S10). All studies revealed that subg. *Poniticae*, subg. *Seriphidium*, subg. *Absinthium*, and subg. *Artemisia* should belong to one large clade, but the phylogenetic relationships among these subgenera remain unresolved (Sanz et al. 2008, Tkach et al. 2008b, Hobbs and Baldwin 2013, Malik et al. 2017, Jiao et al. 2023).

Our results support the division of subg. *Artemisia* into three lineages (Jiao et al. 2023) that are highly correlated with geographic distribution (Figs 2, 4, Supporting Information Figs 12, 13). Among them, *A. selengensis* and *A. anomala* serve as the early lineages of subg. *Artemisia*. The nucleoplasmic conflict observed may be attributed to plastid capture events between these species and *A. palustris*. In the China–Japan region, the overlapping geographical ranges of *A. selengensis*, *A. anomala* and *A. palustris* suggest the potential for hybridization events. Pollen competition is influential in the formation of hybrids (Rieseberg 1995), and it can be hypothesized that the ancestral lineage of Clade I may have served as the maternal parent, and *A. palustris* served as the paternal parent. Continuous backcrossing after hybridization resulted in their progeny possessing the cytoplasm of the maternal parent and the nuclear genes of the paternal parent, thus leading to the plastid similarity of *A. palustris*, *A. selengensis*, and *A. anomala*. Additionally, we detected introgression of Clades III and VII into *A. selengensis* and *A. anomala*, with a more significant introgression from Clade VII into these two species (Fig. 3). This explained why *A. selengensis* and *A. anomala* have a mixture of genetic components of Clades III and VII and clustered together with Clade VII.

N. pectinate, with heterogamous disc florets and functional male-centric florets, was separated from *Artemisia* (Poljakov 1955). However, our results support *N. pectinate* (= *A. pectinate*) as a member of *Artemisia* (Sanz et al. 2008, Jiao et al. 2023). It is noteworthy that *N. pectinate* shows a large variation in its phylogenetic position based on different datasets, and is not explained by plastid capture events (Fig. 2). According to the latest classification under *Artemisia*, *A. pectinate*, and *A. hedinii* were included in the subg. *Pectinata* (Jiao et al. 2023). However, in this study, all the data results showed that they are not closely related, and more molecular and morphological evidence is still needed for the taxonomic delimitation of this species in the subgenus.

In early molecular phylogenetic studies, the relationship between *A. scoparia* and *A. demissa* has not been well explained, and the two species were placed in the same lineage on the basis of genome sizes (Pellicer et al. 2011, 2014). However, our study now provides a clearer picture, indicating that *A. demissa* diverged earlier than *A. scoparia* with robust support (Figs 1, 2). In addition, as the world's widespread taxonomic species, *A. scoparia* and *A. capillaris* have a nested phylogenetic relationship, and both showed polyphy. A similar report has been reported in *A. frigida* (Oyundelger et al. 2022). *A. scoparia* and *A. capillaris* collected from different regions tended to cluster in different branches of the phylogenetic tree (Fig. 4), indicating that they may have undergone adaptive divergence in different regions and formed different lineages.

Phylogenetic results from different datasets showed a high degree of inconsistency, with strong nucleoplasmic conflicts within different lineages. Events such as introgression, ILS, hybridization, and polyploidy may lead to the occurrence of nucleoplasmic conflicts. Major lineages were clearly separated in this study, and introgression occurred mainly within them, but introgression across lineages also occurred (Fig. 3). Highly similar plastomes, introgression, and rapid radiation within the major lineages suggest the presence of plastid capture events between *Artemisia* species. This exacerbates the complexity of species phylogenies within the genus (Kim et al. 2020). At the same time, gene flow across lineages and early reticulation conflicts may suggest ancestral gene flow. Gene flow across lineages is recognized as a common and important process in adaptive radiation and plays a crucial role in facilitating species adaptation to diverse and varied habitats (Berner and Salzburger 2015). ILS and random recombination allowing certain regions of the genome to retain genes acquired from other lineages during ancestral hybridization, this also helps to explain the low resolution and topological inconsistencies in early molecular phylogenetic studies of *Artemisia* (Riggins and Seigler 2012, Hobbs and Baldwin 2013, Shen et al. 2021, Jin et al. 2023). Moreover, some species of *Artemisia* are not reproductively isolated (Vallès et al. 2011). The lack of reproductive isolation, coupled with wind pollination, sets the stage for a plethora of reticulate events, including various types of hybridization, generating multiple and recurrent taxa (García et al. 2011). This has also resulted in ambiguous phylogenetic relationships among taxa and influenced the results of molecular phylogenies of *Artemisia* (Gagnon et al. 2022), as well as explained past studies produced different phylogenetic results based on different molecular data.

Historical biogeography of the major lineages of East Asian *Artemisia*

The presence of *Artemisia* pollen fossils is limited during the Early Miocene, but becomes more prevalent throughout the late Miocene and are notably plentiful during the Pliocene and Pleistocene periods (Wang 2004). This can infer that *Artemisia* probably underwent divergence throughout the early and mid-Miocene epochs, with a significant concentration of diversification events occurring in the late Miocene and Pliocene periods (Sanz et al. 2011). A marked increase in lineage accumulation for *Artemisia* from the Pliocene to Pleistocene period (Supporting Information, Fig. S10) a pattern that corresponds

to the recorded changes in the relative abundance of fossil pollen during the same period. These findings are in alignment with the environmental and climatic fluctuations that occurred in Asia throughout the late Tertiary period (Wang 2004).

The Asian origin of *Artemisia* is undisputed, but the exact region remains controversial. According to Ling (1995) and Wang (2004), *Artemisia*'s predecessor probably originated in north-western Asia, possibly from the mesothermic subarctic or semihumid forest steppe habitats, based on evidence from paleobotanical and pollen fossils. Tkach *et al.* (2008) used ITS and ETS to reveal the Arctic origins of several lineages under the modern *Artemisia*. A recent historical biogeographic analysis of the subg. *Seriphidium* suggests that *Artemisia* originated in the arid-semi-arid mid-latitudes of Asia in the Late Eocene (Malik *et al.* 2017), but this study categorizes both the Hengduang-Himalaya and the QTP as Tethyan. This contradicts the delineation of the East Asian flora (Wu *et al.* 2011, Chen *et al.* 2018a). This study redefined the distribution region of *Artemisia* (Fig. 4) by referring to studies on East Asian flora (Wu *et al.* 2011, Chen *et al.* 2018a) and world *Artemisia* flora (Ling 1995). The results indicated the modern East Asian *Artemisia* was probably diversified *in situ* in the Early Oligocene. The early divergence of *Artemisia* coincided with the strengthening of the Asian monsoon and climate cooling during the Oligocene–Miocene (Deng *et al.* 2019). The major lineages diverged in the Early Miocene (Sanz *et al.* 2011), a period in which orogenic movements and climatic reorganization in East Asia combined to promote species dispersal and diversification (Li *et al.* 2021, 2023).

The subg. *Artemisia* in this study can be divided into three lineages: Clade VI is distributed only within China–Japan and Hengduan–Himalaya; Clade VII-1 gathers most of the taxa that are distributed only in the Hengduan–Himalaya; and Clade VII-2 gathers a larger number of species that are widely distributed with a distinct high-latitude component (Fig. 4). Among them, Clade VII-1 and Clade VII-2 occurred mainly in the mid-Miocene (13.78–8.45 Mya), corresponding to the global climatic transition and the ongoing geotectonic events in the Himalayas (Shevenell *et al.* 2004, Steinthorsdottir *et al.* 2021, Sun *et al.* 2023). The complex topography, heterogeneous habitats, and dispersal barriers resulting from continued orogeny, lower temperatures, and the intensification of the Asian monsoon climate during this period may have facilitated the occurrence of parallel speciation in both clades (Zhisheng *et al.* 2001). Parallel species formation events, similar to those observed in *Artemisia*, have been reported in a recent phylogenetic study of *Rhododendron* L. (Mo *et al.* 2022).

The subg. *Poniticae* and subg. *Absinthium* diversifies at almost the same time, and both exhibit a more prominent Tethyan element to their diversification. The distinction lies in the fact that subg. *Poniticae* tends to diversify at high latitudes, whereas the low-latitude ancestral range component is more pronounced in subg. *Absinthium* during diversification (Fig. 4, Supporting Information Figs S12, S13). This phase coincided with the mid-Miocene period of climatic suitability, when suitable climate and uplift of the Mongolian Plateau may have influenced the early divergence of these two subgenera (Xiang *et al.* 2021). The following QTP–Himalaya uplift led to asymmetric climatic evolution in the Asian monsoon and inland arid zones (Lewis *et al.*

2008, Ao *et al.* 2023), which could lead to further divergence and adaptive evolution of the two lineages. In addition, the relatively strong uplift that occurred after the Cenozoic in places such as the Ural Mountains and the Mongolian Plateau (Wang *et al.* 2019) have further exacerbated the adaptive evolution of subg. *Absinthium*.

Subg. *Dracunculus* has ~80 species and is distributed mainly in steppes and arid to semi-arid regions throughout the Northern Hemisphere, but it is particularly diverse in East Asia (Pellicer *et al.* 2011). In this study, phylogenetic results based on nuclear data showed that this subgenus is divided into two lineages (Pellicer *et al.* 2011, Jiao *et al.* 2023). Early Miocene cooling and monsoon climate (Zachos *et al.* 2001, Cliff *et al.* 2002) together contributed to the diversification of the subg. *Dracunculus*. In comparison to Clade V-1, species in Clade V-2 exhibit a larger *ycf1* pseudogene (*Ψycf1*) deletion (Fig. 4, Supporting Information Table S2), which may be related to the recent rapid diversification (Malik *et al.* 2017).

This study confirms that subg. *Seriphidium* genealogy belongs to *Artemisia* (Jiao *et al.* 2023), with an intimate phylogenetic relationship to subg. *Absinthium*, in general agreement with previous findings (Malik *et al.* 2017, Jin *et al.* 2023). We verified that the subg. *Seriphidium* is the youngest taxon in *Artemisia* (Riggins and Seigler 2012, Malik *et al.* 2017). The core clade of the subg. *Seriphidium* originated around the late Miocene–early Pliocene and underwent a major diversification transition (Malik *et al.* 2017). Our estimate of a slightly later divergence than the previous study (Malik *et al.* 2017) may be related to sampling. According to Hobbs (2013) subgeneric taxonomic classification, we detect differences in the plastid structure of subg. *Seriphidium* between species that belong to Clade III and those that clustered with subg. *Absinthium*: the latter contains the *trnT-GGU* gene while the former does not (Supporting Information, Table S3). This gene is one of the highly variable regions in the plastomes of subg. *Seriphidium* (Jin *et al.* 2023). It has been proven that in the subfamily Asteroideae, pseudogenization of *trnT-GGU* is connected to insertion inside the 5' acceptor stem rather than habit, habitat, or geographical distribution (Abdullah *et al.* 2021).

The early divergent lineages of East Asian *Artemisia*

Although introgression mainly occurs within the major lineages, there are still some introgressions across the lineages. However, *A. pectinate*, *A. palustris*, *A. keiskeana*, and *A. hedinii* show more complex genetic components than most species of *Artemisia*, but we did not detect significant introgression in these species. Introgression from multiple lineages and species could explain the complex genetic components of *A. pectinate*, *A. palustris*, and *A. hedinii* (Fig. 3). By contrast, there was no evidence of introgression between the main lineages and several species in *A. keiskeana*. All the available data suggest that *A. keiskeana* represents an ancient branch of East Asian *Artemisia*. These species tended to be located at the base of the major lineages (Clade V, Clade VII + Clade VIII + Clade IX) in the phylogeny based on different datasets (Fig. 2), which is generally consistent with the results of the previous studies (Pellicer *et al.* 2011, Malik *et al.* 2017, Jiao *et al.* 2023). These four species diverged in the Oligocene–Miocene, predating the diversification of the major lineages under *Artemisia* (Fig. 4). Therefore, we speculate that

species such as *A. pectinate*, *A. palustris*, *A. keiskeana*, and *A. hedinii*, which have mixed genetic components (Supporting Information, Fig. S9), may be the early diverging lineages of major lineages of modern *Artemisia* in East Asia. The presence of a mixed genetic architecture has the potential to serve as an optimal foundation for the quick process of speciation, even in the absence of geographic isolation. The reason for their potential to more effectively promote the accumulation of linkage disequilibrium in the presence of gene flow is greater in comparison to either a limited number of crucial loci or highly polygenic topologies (Nosil *et al.* 2009). The occurrence of secondary contact among the early diverging lineages may have played a role in promoting the rapid and repeated diversification of species that exhibit similar morphological features under *Artemisia*. This also can explain the differences in the phylogenetic positions of these species in different datasets. However, it can also impose constraints on the development of phenotypic divergence, which could otherwise be readily achieved through the recombination of existing alleles (Marques *et al.* 2019, De-Kayne *et al.* 2022).

Nevertheless, we cannot rule out the possibility that the early diverging lineages actually functioned as ‘fused species’ with mixed genetic architecture, supplying the gene pool with genomic fractions adapted to heterogeneous environments and having the potential to diverge into multiple species in specific circumstances (Edelman *et al.* 2019, Yu *et al.* 2023). The Early Miocene global cooling event, which lasted ~23 Mya and exceeded 400 kya in duration (Miao *et al.* 2013), is a significant period during which these early diverging lineages underwent divergence. This prolonged cooling event introduces uncertainty regarding the extinction of the predecessors of the early diverging lineages throughout the extensive periods of glaciation. The ‘fused taxa’ with mixed genetic architecture successfully migrated via refugia and subsequently underwent growth and diversification, leading to the establishment of the now thriving *Artemisia* taxa.

CONCLUSION

In this study, we applied genome skimming techniques to obtain multiple molecular datasets, established robust phylogenetic relationships of *Artemisia* and its subgenera in East Asia, and provided the first insight into the historical biogeography. East Asian *Artemisia* may have originated *in situ* in the Early Oligocene. The diversification of *Artemisia* and its subgenera were impacted by historical geographic events and climatic fluctuations. *A. pectinate*, *A. palustris*, *A. keiskeana*, and *A. hedinii* are regarded as the early diverging lineages of modern East Asian *Artemisia*, possessing mixed genetic components. Introgression occurs mainly within major lineages, and reticulation events affect and obscure deep phylogeny. Our study reveals the complex phylogenetic relationships and evolutionary history of East Asian *Artemisia* and its subgenera. Further investigation is required through the implementation of more comprehensive sampling and in-depth studies to gain a more profound understanding of this intricate taxonomic group.

SUPPLEMENTARY DATA

Supplementary data is available at *Botanical Journal of the Linnean Society* online.

CONFLICT OF INTEREST

None declared.

FUNDING

This research was partially supported by the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (2019QZKK0502), Qinghai Provincial Science and Technology Major Project (2023-SF-AS), the Biological Resources Programme of Chinese Academy of Sciences (KFJ-BRP-017–101), CAS—Qinghai on Sanjiangyuan National Park (LHZZ-2021–04), and CAS ‘Light of West China’ Program (2024).

DATA AVAILABILITY

The data on which this article is based are available from the GenBank nucleotide database at <https://www.ncbi.nlm.nih.gov/genbank/>, and the accession number is in Table S3 in the Supporting Information.

REFERENCES

- Abad MJ, Bedoya LM, Apaza L *et al.* The *Artemisia* L. genus: a review of bioactive essential oils. *Molecules* 2012;17:2542–66. <https://doi.org/10.3390/molecules17032542>
- Abdullah, Mehmood F, Heidari P *et al.* Pseudogenization of the chloroplast threonine (trnT-GGU) gene in the sunflower family (Asteraceae). *Scientific Reports* 2021;11:21122. <https://doi.org/10.1038/s41598-021-00510-4>
- Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 2009;19:1655–64. <https://doi.org/10.1101/gr.094052.109>
- Ao H, Liebrand D, Dekkers MJ *et al.* Eccentricity-paced monsoon variability on the northeastern Tibetan Plateau in the Late Oligocene high CO₂ world. *Science Advances* 2023;7:eabk2318.
- Berner D, Salzburger W. The genomics of organismal diversification illuminated by adaptive radiations. *Trends in Genetics* 2015;31:491–9. <https://doi.org/10.1016/j.tig.2015.07.002>
- Chen S, Zhou Y, Chen Y *et al.* fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018b;34:i884–90. <https://doi.org/10.1093/bioinformatics/bty560>
- Chen YS, Deng T, Zhou Z *et al.* Is the East Asian flora ancient or not? *National Science Review* 2018a;5:920–32. <https://doi.org/10.1093/nsr/nwx156>
- Clift P, Lee JJ, Clark MK *et al.* Erosional response of South China to arc rifting and monsoonal strengthening; a record from the South China Sea. *Marine Geology* 2002;184:207–26. [https://doi.org/10.1016/S0025-3227\(01\)00301-2](https://doi.org/10.1016/S0025-3227(01)00301-2)
- Danecek P, Auton A, Abecasis G *et al.*; 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics* 2011;27:2156–8. <https://doi.org/10.1093/bioinformatics/btr330>
- Degnan JH, Rosenberg NA. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* 2009;24:332–40. <https://doi.org/10.1016/j.tree.2009.01.009>
- De-Kayne R, Selz OM, Marques DA *et al.* Genomic architecture of adaptive radiation and hybridization in Alpine whitefish. *Nature Communications* 2022;13:4479. <https://doi.org/10.1038/s41467-022-32181-8>
- Deng T, Wu F, Wang S *et al.* Significant shift in the terrestrial ecosystem at the Paleogene/Neogene boundary in the Tibetan Plateau. *Chinese Science Bulletin* 2019;64:2894–906.
- De Tournefort JP. *Institutiones rei herbariae*: Vol. 1. *Typographia regia* 1719;1:i57–68.
- Doyle JJ. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bulletin* 1987;19:11–5.
- Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 2007;7:214. <https://doi.org/10.1186/1471-2148-7-214>

