BMC Plant Biology

RESEARCH Open Access

Revisiting *Phryma leptostachya* L.: phylogenetic relationships and biogeographical patterns from complete plastome

Yeseul Kim[1], Sumin Jeong[1], Inkyu Park[1*] and Hye-Kyoung Moon[2*]

Abstract

Background *Phryma leptostachya* L. is a notable example of a species with a disjunct distribution, found in both East Asia and Eastern North America. Despite the striking morphological similarities between these geographically isolated populations, molecular evidence suggests that they may have diverged sufficiently to be considered distinct taxa.

Results To clarify this, we analyzed the plastomes of *P. leptostachya* from Korea, Russia, and the USA. Their sizes ranged from 152,974 to 153,325 bp, each containing 113 genes. Differences were observed in the boundaries between large single copy (LSC)/IRa and IRb/LSC. In *P. leptostachya*_USA, the *rps19* gene extended 30–31 bp into the IRa, and the *rpl2* gene contracted 51–53 bp at the IRa/b compared to those of *P. leptostachya*_Korea and *P. leptostachya*_Russia, suggesting that expansion of the inverted repeat (IR) region occurred in *P. leptostachya*_USA. Regions such as *psbZ-trnG*, *ccsA*-ndhD, *petA-psbJ*, and *psbC-trnS* were identified as hotspots with sequence differences in the plastome, indicating differences among *P. leptostachya* variants. Phylogenetic analysis showed that *P. leptostachya* from Korea and Russia formed monophyletic groups, while the variety from the USA was paraphyletic. The divergence of *P. leptostachya*_USA occurred during the Pliocene, about 5.25 million years ago (MYA), whereas the split between *P. leptostachya*_Korea and *P. leptostachya*_Russia is estimated to have occurred approximately 0.87 MYA during the Pleistocene. The results also reveal that the family Phrymaceae underwent multiple dispersal and vicariance events from North America to East Asia, offering key insights into the phylogenetic relationships between *P. leptostachya* populations from Korea, Russia, and the USA. Based on the evidence, it is likely that *P. leptostachya* originated in North America and later migrated to East Asia via the Russian Far East and the Bering Land Bridge.

Conclusions In conclusion, our study demonstrates clear molecular differences among *P. leptostachya* populations from various geographic locations, suggesting that these populations should be recognized as distinct species rather than conspecifics.

*Correspondence:
Inkyu Park
pik6885@cwnu.ac.kr
Hye-Kyoung Moon
hkmoon@khu.ac.kr

Full list of author information is available at the end of the article

Keywords *Phryma leptostachya*, Evolutionary divergence, Phylogenetic relationships, Intercontinental distribution, Genetic variability, Biogeography

Background

Phryma leptostachya L. is a temperate perennial plant belonging to the family Phrymaceae [1, 2]. The genus *Phryma* was initially classified within Verbenaceae due to its morphological similarities with other members of that family [1, 3, 4]. However, it was later reclassified as a new monotypic family, Phrymaceae, based on its distinctive pseudomonomerous gynoecium, characterized by a two-carpellate structure with one carpel being developmentally reduced [5, 6]. Subsequent phylogenetic analyses revealed a significant expansion of the family after the transfer of genera from Scrophulariaceae [7–9]. Currently, Phrymaceae consists of approximately 14 genera and 150 species, including *Cyrtandromoea*, *Diplacus*, *Erythranthe*, *Mimulus*, and *Phryma* [9, 10].

Studies on divergence time have predominantly focused on disjunct taxa in temperate regions, with previous research conducted on *Liriodendron* [11], *Liquidambar* [12], *Campsis* [13], and *Symplocarpus* [14]. *P. leptostachya* is a classic example of the eastern Asia and eastern North American disjunction pattern. This disjunct distribution pattern is commonly observed in the Northern Hemisphere, reflecting a classic biogeographic pattern shaped by historical factors such as the Bering Land Bridge, the North Atlantic Land Bridge, and global climate changes during the Tertiary and Quaternary periods [15–18]. Two varieties are traditionally recognized: var. *leptostachya* in North America and var. *asiatica* H. Hara in eastern Asia [1]. Although these varieties share highly similar morphological characteristics, they differ in traits such as calyx tube length and corolla lip shape, and genetic analysis reveals distinct molecular divergence between them [2]. Research using allozyme data has shown low genetic identity between them, indicating a significant period of divergence [19]. Furthermore, both allozyme data and ITS sequences suggest lineage splitting [20]. Despite morphological similarities [2], intercontinental populations of *Phryma* exhibit significant molecular divergence according to molecular data [15]. However, previous studies have been limited to specific DNA regions, including nuclear ribosomal ITS, chloroplast *rps16*, and *trnL-F*.

The plastome, maternally inherited and characterized by infrequent recombination and a slower mutation rate compared to the nuclear genome, serves as a valuable tool in evolutionary biology research for reconstructing relationships between plants [21, 22]. Advances in sequencing technologies have facilitated more accurate biogeographic mapping and improved the usefulness of the plastome, which is highly conserved among plant species [23–25]. Two plastomes of *Phryma leptostachya* (subsp. *asiatica*), a species within the Phrymaceae, are available in the NCBI database: NC_042727.1 and MT948145.1. These plastomes are extensively utilized for phylogenetic and biogeographic analyses, providing critical insights into the evolutionary history of angiosperms [26–33]. Using the plastome rather than specific DNA fragments may provide a more comprehensive understanding of the molecular systematics of *Phryma*, thereby enhancing our understanding of its evolutionary history, taxonomic relationships, and species distribution evolution [34, 35].

P. leptostachya is a discontinuously distributed intercontinental species, with populations in East Asia and North America recognized as either a single species or distinct taxa, depending on interpretation. Previous studies have recognized *P. leptostachya* as comprising two varieties; however, these studies have not sufficiently resolved this taxonomic issue. In this study, we used plastomes instead of partial sequences and included samples from East Asia, North America, and the Russian Far East. The primary goal of our study was to investigate geographically distinct *P. leptostachya* and conduct a high-resolution analysis of intercontinental migration patterns. To achieve this, we (1) sequenced and compared the plastomes collected from Korea, the Russian Far East, and the USA to identify genomic differences; (2) performed phylogenetic analyses, estimated divergence times, and inferred evolutionary relationships; and (3) reconstructed the biogeographic history of *Phryma* to confirm and elucidate intercontinental migration patterns.

Methods

Plant materials

Fresh samples of *Phryma leptostachya* were collected directly from Gayasan Mt., Hapcheon-gun, Gyeongsangnam-do, South Korea (KHUM20140008: 35.800323°N, 128.098229°E) by H.K. Moon. Additionally, samples from Vladivostok, Primorsky Krai, Russia (KHUJ20140001: 43.177034°N, 131.946176°E), and Athens County, Ohio, USA (KHUC20141453: 39.344397°N, 82.087248°W) were kindly provided by collaborators. These voucher specimens were deposited in the herbarium of Kyung Hee University. For barcoding analysis, dried samples were obtained from herbarium specimens at the Meise Botanic Garden in Belgium (BR), the National Institute of Biological Resources in South Korea (KB), and the Naturalis Biodiversity Center in the Netherlands (L), with

appropriate permissions. The detailed collection information is listed in Tables S1 and S2.

Genome sequencing, assembly, and annotation

The total DNA of *P. leptostachya* was extracted using a modified cetyltrimethylammonium bromide method [36]. After evaluating and confirming the quality of the DNA samples, we prepared genome libraries using the TruSeq Nano DNA Kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. Sequencing was then performed on the MiSeq platform to generate 22–24 Gb of paired-end reads. To improve the data quality, we used Trimmomatic version 0.39 with Phred33 encoding [37]. To optimize read quality, we applied options for adapter removal, low-quality base trimming, sliding window trimming, and removal of reads shorter than 36 bases. Clean reads were assembled using the GetOrganelle program [38] with the reference sequence for *P. leptostachya* subsp. *asiatica* (MT948145.1; *P. leptostachya*_Korea2). To determine the initial position and orientation of the plastome assembly sequence, as well as the boundaries between regions, a self-blast was performed. The plastomes of *P. leptostachya* were annotated using the GeSeq tool [39] to predict protein-coding sequences (CDSs), tRNA, and rRNA genes. Final manual corrections were performed using Geneious software [40]. Subsequently, the accuracy of the assembled plastomes was verified by mapping the plastome data from *P. leptostachya* to their respective assembled sequences using Burrows-Wheeler Aligner (BWA) [41] and Sequence Alignment/Map (SAM) tools [42]. Circular genome maps were generated using OGDRAW [43]. The complete plastome sequences and annotations of *P. leptostachya* were submitted to GenBank (PQ443460-PQ443462).

Codon usage and comparative analysis

Relative Synonymous Codon Usage (RSCU) represents the ratio of observed codon usage frequency to the expected unbiased frequency, indicating codon preference. An RSCU < 1.00 signifies a lower frequency, while RSCU > 1.00 indicates a higher frequency. We analyzed the RSCU and GC content using MEGA11 software [44] and visualized codon usage patterns using Heatmapper [45]. Sequence and structural differences in the plastomes were visualized using the Shuffle-LAGAN mode of the mVISTA program [46], with *P. leptostachya*_Korea1 as the reference genome. To calculate nucleotide diversity, sequences of *P. leptostachya*_Korea1 were aligned with those from Korea2, Russia, and the USA separately using Geneious software [40]. Subsequently, CDS were extracted, and nucleotide diversity was calculated using DnaSP version 6 [47].

Repeat analysis

Simple Sequence Repeats (SSRs) were identified using MISA software [48] with the following parameters: 10 mononucleotides, 5 dinucleotides, 4 trinucleotides, 3 tetranucleotides, 3 pentanucleotides, and 3 hexanucleotides. REPuter software [49] was used to detect repeats with a minimum length of 30 bp, a repeat identity of at least 90%, and a Hamming distance of 3. The repeats were categorized as forward (F), reverse (R), palindromic (P), and complementary (C). For tandem repeat analysis, we used the Tandem Repeats Finder [50], with a minimum alignment score of 50 and a maximum period size of 500. The alignment parameters were set as follows: a match score of 2, a mismatch penalty of 7, and an indel penalty of 7. Additionally, we considered only repeats with an identity of 90% or greater.

Selective pressure analysis

The dN/dS ratio ($\omega = dN/dS$) of protein CDSs was calculated using PAML package v.4.9 [51], with reference to *Erythranthe dentiloba*. An $\omega > 1$ indicates positive selection, $\omega = 1$ indicates neutral evolution, and $\omega < 1$ indicates purifying selection. The yn00 module was selected to estimate nonsynonymous substitution rate (dN) and synonymous substitution rate (dS) using the following parameters: 'verbose = 0, icode = 0, weighting = 0, commonf3 × 4 = 0, ndata = 1.' Subsequently, boxplots were generated using R software to visualize the results.

Phylogenetic analysis

We downloaded the plastomes of 17 species, including outgroups (NC_037506.1, *Lancea hirsuta*; NC_037693.1, *L. tibetica*; NC_056337.1, *Dodartia orientalis*), from NCBI based on the APG IV system [52] (Table S3). We then used MAFFT version 7.388 to align their positions [53], and 78 CDS were extracted in alphabetical order using Geneious software [40]. To enhance alignment accuracy, ambiguously aligned regions and gaps were identified and removed using GBlock version 5 [54]. After refining the alignment, we selected the optimal model for the aligned sequences using the Akaike Information Criterion provided by JModelTest version 2.1.10 [55], as detailed in Table S4. The GTR + I + G model was selected for phylogenetic analysis. Maximum likelihood (ML) trees were constructed using 1,000 bootstrap replications performed using MEGA 11 [44]. For the Bayesian Inference (BI) analysis, we used MrBayes version 3.2.2 [56] included in Geneious software [40] to obtain a robust tree. A GTR model with four gamma categories was set, and the MCMC method was run for 5,000,000 generations. The tree was sampled at intervals of 5,000 generations, with the initial 25% of the samples discarded as burn-ins.

Divergence time estimation

Divergence time was estimated using BEAST version 2.7.6 [57] with 78 protein-coding genes (PCGs) from 20 plants. A GTR replacement model was designed with gamma distributions across four ratio categories and estimated using a Yule Prior. For divergence time analysis, two calibration points were established: (1) Lamiales segmentation at 85.12–89.91 million years ago (MYA) and (2) Phrymaceae segmentation at 27.86–52.53 MYA [15]. MCMC chains were run for 50 million generations, and sampled every 10,000 generations. The effective sample size (ESS) was evaluated using Tracer version 1.7.2 [58], with all parameters having ESS values greater than 200, indicating sufficient sampling. The first 10% of trees were discarded as burn-ins, and maximum clade credibility trees were generated using TreeAnnotator version 2.7.6 [57]. These were visualized using Figtree version 1.4 [59], and the results were determined using the mean height and 95% high posterior densities (HPDs).

Reconstructing ancestral states

Biogeographic data for the Phrymaceae, Paulowniaceae, and Mazaceae outgroup species used in this study were obtained from the World Online database (https://powo.science.kew.org) and previously published papers [60–65]. Statistical dispersal variance analysis (S-DIVA) was implemented to infer the ancestry distribution using RASP version 4.4 software [66]. BEAST and consensus trees constructed from a dataset consisting of 78 PCGs were used as the input data.

Haplotype network analysis

To further investigate the intercontinental differences, we used the general-purpose DNA barcodes ITS2 and *matK* for 21 *P. leptostachya* samples. *MatK* was redesigned using Primer3Plus [67]. DNA extracted from 21 samples was amplified via PCR using 10 ng of template DNA. The PCR reaction was carried out using a 20 µL mixture containing 10 pmol of primer in a PCR system. Amplification parameters were as follows: initial denaturation at 95 °C for 2 min; 35 cycles at 95 °C for 50 s, 58 °C for 50 s, and 72 °C for 50 s; final extension at 72 °C for 5 min. The PCR products were separated on a 2% agarose gel at 150 V for 40 min. The germplasms of the 21 *P. leptostachya* samples are listed in Table S2, and the primer sequences are provided in Table S5.

The ITS2 and *matK* sequences were used for haplotype analyses (Fig. S1). Based on the aligned sequences, a haplotype data file was generated using DnaSP version 6 [47]. Sequence sets were assigned as follows: (A) East Asia, (B) North America, and (C) the Russian Far East. The haplotype network was constructed and plotted using the median joining method in Popart version 1.7 [68].

Results

Characterization of *Phryma leptostachya* plastome

The plastomes of *P. leptostachya* were analyzed to generate 22–24 Gb of data (Tables S6 and S7). *P. leptostachya*_Korea1, *P. leptostachya*_Russia, and *P. leptostachya*_USA had coverages of 3,349 ×, 1,241 ×, and 2,771 ×, respectively (Table S8). Read mapping of the plastome datasets to each genome revealed that all were of high quality (Fig. S2). The assembled plastomes of *P. leptostachya* had a typical quadripartite structure consisting of a large single copy (LSC), a small single copy (SSC), and two inverted repeat regions (IRa and IRb) (Fig. 1 and S3). Plastome size varied from 152,974 to 153,325 bp. The LSC regions ranged from 84,728 to 84,997 bp, the SSC regions from 17,500 to 17,544 bp, and the inverted repeat (IR) regions from 25,373 to 25,393 bp (Table 1). The overall GC content was 37.72–37.74%, with LSC regions at approximately 36%, SSC regions around 32%, and IR regions at about 43%. They contained 113 genes, including 79 PCGs, 30 tRNAs, and 4 rRNAs. Eighteen genes in the *P.*

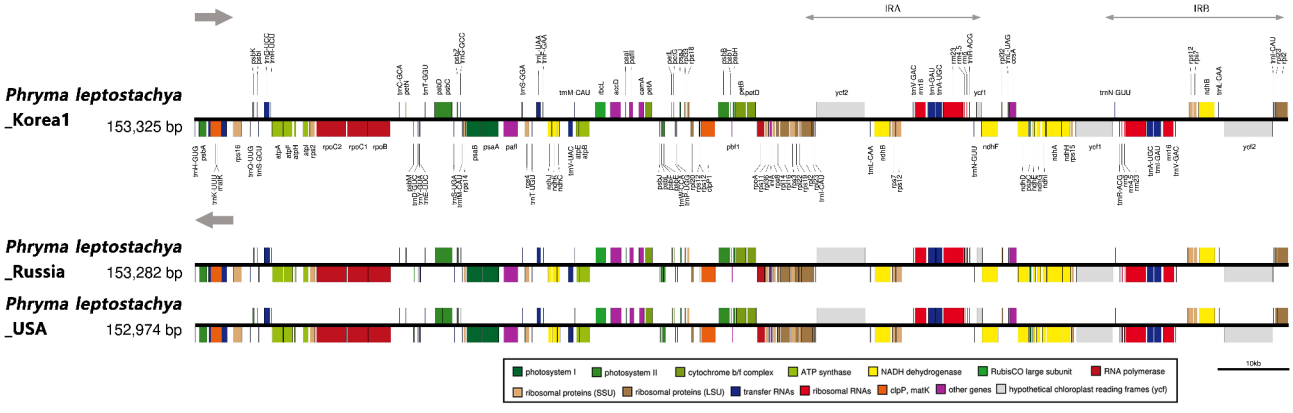

Fig. 1 Gene map of the complete plastome of *P. leptostachya* represented in a linearized form. Genes are transcribed from left to right. The arrow above the gene indicates the forward direction (left to right), while the arrow below the gene indicates the reverse direction (right to left). The gray line indicates IR regions

Table 1 Features of *Phryma leptostachya* plastomes. The *P. leptostachya* Korea sample presented in the table corresponds to *P. leptostachya*_Korea1

Species	*P. leptostachya*	*P. leptostachya*	*P. leptostachya*
Origin	Korea	Russia	USA
Accession number	PQ443460	PQ443461	PQ443462
Total plastome size (bp)	153,325	153,282	152,974
Large single copy (LSC) region (bp)	84,997	84,952	84,728
Inverted repeat (IR) region (bp)	25,393	25,393	25,373
Small single copy (SSC) region (bp)	17,542	17,544	17,500
Total number of genes (unique)	113	113	113
Protein-coding gene (unique)	79	79	79
rRNA (unique)	4	4	4
tRNA (unique)	30	30	30
GC content (%)	37.72	37.72	37.74
LSC (%)	35.73	35.73	35.75
IR (%)	43.14	43.13	43.16
SSC (%)	31.66	31.66	31.65

leptostachya plastomes contained introns, including 16 genes with single introns. Of these, two (*pafI* and *clpP1*) contained duplicate introns (Tables S9 and S10).

We estimated codon usage and anticodon recognition patterns of the plastome based on PCGs (Fig. S4). The results showed that leucine was the most abundant amino acid, whereas cysteine was the least abundant. Except for methionine and tryptophan, which were encoded by only one codon, the remaining amino acids contained 2–6 codons. The codon distribution was visualized as a heat map (Fig. S5), where colors indicate codon bias, in which green represents RSCU > 1, and red represents RSCU < 1. We observed that the RSCU values of codons ending at the A/T position in the third position were high.

Repetitive DNA sequences

We performed repeat analyses of the four *P. leptostachya*. In *P. leptostachya*_Russia, we detected 20 forward and 30 palindromic repeat types, whereas the other three samples exhibited 21 forward and 29 palindromic repeat types (Fig. 2A). The SSR analysis identified mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeat motifs, with mononucleotides being the most abundant, followed by tetranucleotides (Fig. 2B). The number of SSRs ranged from 36 to 44, with many primarily distributed in the intergenic spacer (IGS) and LSC regions (Fig. 2B). *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2 had identical SSR numbers and distributions, whereas *P. leptostachya*_USA had the lowest number of SSRs and was absent from the IR region. No pentanucleotides were detected in *P. leptostachya*_Russia. Tandem repeat analysis showed identical repeat numbers and distributions in *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2, and *P. leptostachya*_Russia and *P. leptostachya*_USA exhibited the same pattern (Fig. 2C). Most tandem repeats were distributed in the IGS and LSC regions, with lengths primarily between 30 and 71 bp (Fig. 2D). Although IGS regions are present throughout the plastome, they are most abundant in the LSC region due to its larger size and gene density. *P. leptostachya*_USA had the highest number of 31–50 bp tandem repeats and the fewest repeats of 30 bp or shorter. The total number of repeat sequences was highest in *P. leptostachya*_Korea1 (3,018 sequences) and lowest in *P. leptostachya*_USA (2,880 sequences). Although the number of tandem repeats was lower than that of the SSRs, the tandem repeats were longer (Fig. 2E). Overall, the percentage of palindromic repeats was highest, followed by forward repeats, tandem repeats, and SSRs, showing similar distribution patterns across *P. leptostachya* (Fig. 2F).

Comparative analysis of four *Phryma leptostachya* plastomes

The contraction and expansion of the IR region are related to the size of the plastome and can indicate an evolutionary relationship between species (Fig. S6). The plastome lengths of *P. leptostachya* ranged from 152,974 to 153,325 bp, with the LSC region ranging from 84,728 to 85,000 bp, SSC region from 17,500 to 17,544 bp, and IR region from 25,373 to 25,393 bp. The IR lengths of *P. leptostachya*_Korea1 and *P. leptostachya*_Russia were the longest at 25,393 bp, whereas the IR length of *P. leptostachya*_USA was the shortest at 25,373 bp. *rps19* and *trnH-GUG* were located in the LSC region, and *rpl2* was found in both the IRa and IRb regions. The SSC region contained the *ndhF* gene and was 35 bp in length. The IRb region, containing the *trnN-GUU* and *ycf1* genes, was located at the boundary between the IRa/SSC and SSC/IRb, crossing the SSC/IR region. The *rps19* gene was 279 bp across all *P. leptostachya* plastomes. In *P. leptostachya*_USA, *rps19* spanned the boundary between the LSC and IRa, extending into the IRa by 42 bp, whereas in *P. leptostachya*, *rps19* was present only in the LSC region. The *ycf1* gene, a pseudogene, was found in two *Diplacus*, two *Erythrnathe*, one *Mimulus*, and one *Paulownia* species.

Sequence identity was determined using the mVISTA program, based on *Phryma leptostachya*_Korea1. The plastome structure of *P. leptostachya* was conserved, with the gene regions being more conserved than the IGS regions (Fig. 3). The IR region was more conserved than the LSC and SSC regions, and many divergent regions were observed in *P. leptostachya*_USA compared to other *P. leptostachya* varieties. Nucleotide diversity (Pi) values

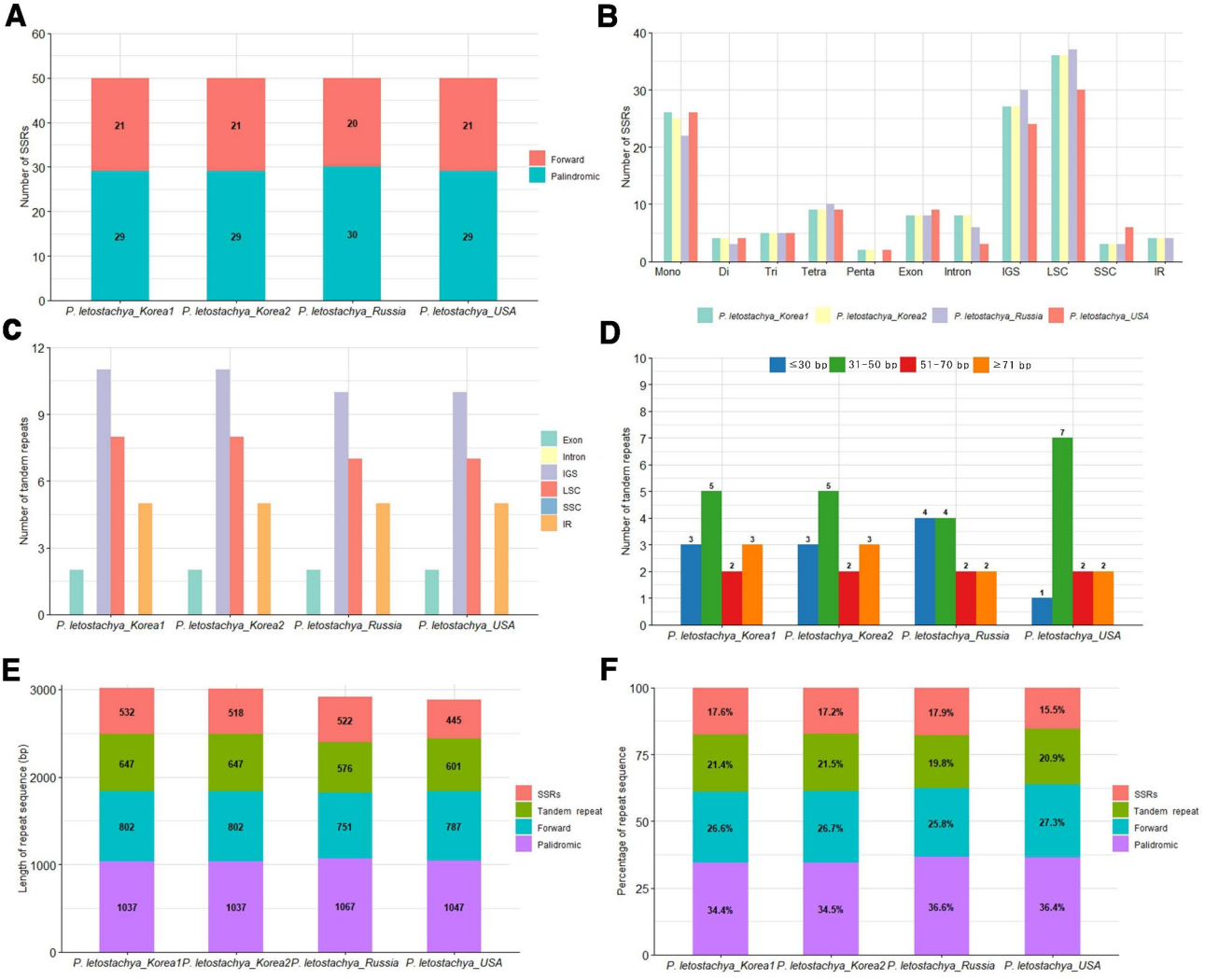

Fig. 2 Distribution of repeats sequence of the four plastomes of *P. leptostachya*. **A** Number of forward, palindromic repeats. **B** Distribution of SSR motif types and number of SSRs. **C** Number of tandem repeats in exons, introns, intergenic spacers (IGS), and genomic regions. **D** Distribution of tandem repeats. **E** Total length of repeat sequences. **F** Percentage of total repeat sequences

were analyzed by comparing *P. leptostachya*_Korea2, *P. leptostachya*_Russia, and *P. leptostachya*_USA with *P. leptostachya*_Korea1 (Fig. 4). The IR region was more conserved than the other regions, and the Pi value in the IGS region was high. Nucleotide diversity between *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2 was nearly absent, with an average value of 0.00027. Between *P. leptostachya*_Korea1 and *P. leptostachya*_Russia, *psbZ-trnG* (0.01408) showed high Pi values. The nucleotide diversity between *P. leptostachya*_Korea1 and *P. leptostachya*_USA showed an average value of 0.00762. The *ccsA-ndhD* region showed the highest diversity (0.04242), followed by *petA-psbJ* (0.02275) and *psbC-trnS* (0.02137).

We performed a selection pressure analysis using *E. dentiloba* as a reference to investigate the rate of genetic evolution (Figs. 5 and S7). In the dN/dS ratio (ω), $\omega = 1$ indicates neutral evolution, $\omega > 1$ indicates positive selection, and $\omega < 1$ indicates purifying selection. Here, dN represents the rate of nonsynonymous mutations that alter the amino acid sequence, whereas dS represents the rate of synonymous mutations that do not alter the amino acid sequence. In the *P. leptostachya* plastome, the overall dN and dS values ranged from 0 to 0.1158 and 0–0.477, respectively. The average dN/dS ratios for *P. leptostachya*_Korea1, *P. leptostachya*_Korea2, *P. leptostachya*_Russia, and *P. leptostachya*_USA were 0.1868, 0.1865, 0.1855, and 0.183, respectively. Most of the 78 PCGs were conserved, indicating a purifying selection. The *psbH* gene, which is involved in photosynthesis, had a value of $\omega > 1$, suggesting that it may be subject to positive selection. This gene was identical in all *P. leptostachya* variants, with a value of 1.0089.

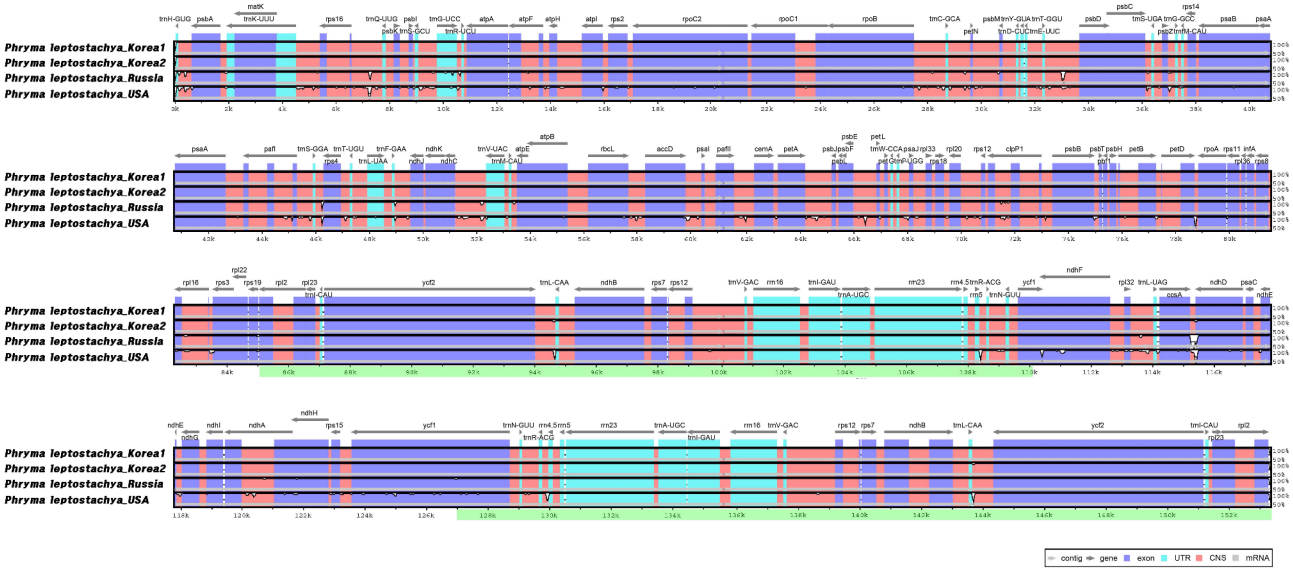

Fig. 3 Comparison of the complete plastomes of four *Phryma leptostachya* using mVISTA, with *P. leptostachya*_Korea1 as the reference. Blue blocks represent conserved genes, sky-blue blocks indicate transfer RNA (tRNA) and ribosomal RNA (rRNA), and red blocks denote conserved non-coding sequences (CNS). Green sections correspond to IR regions, while white regions indicate sequence variation among the four *P. leptostachya* plastomes

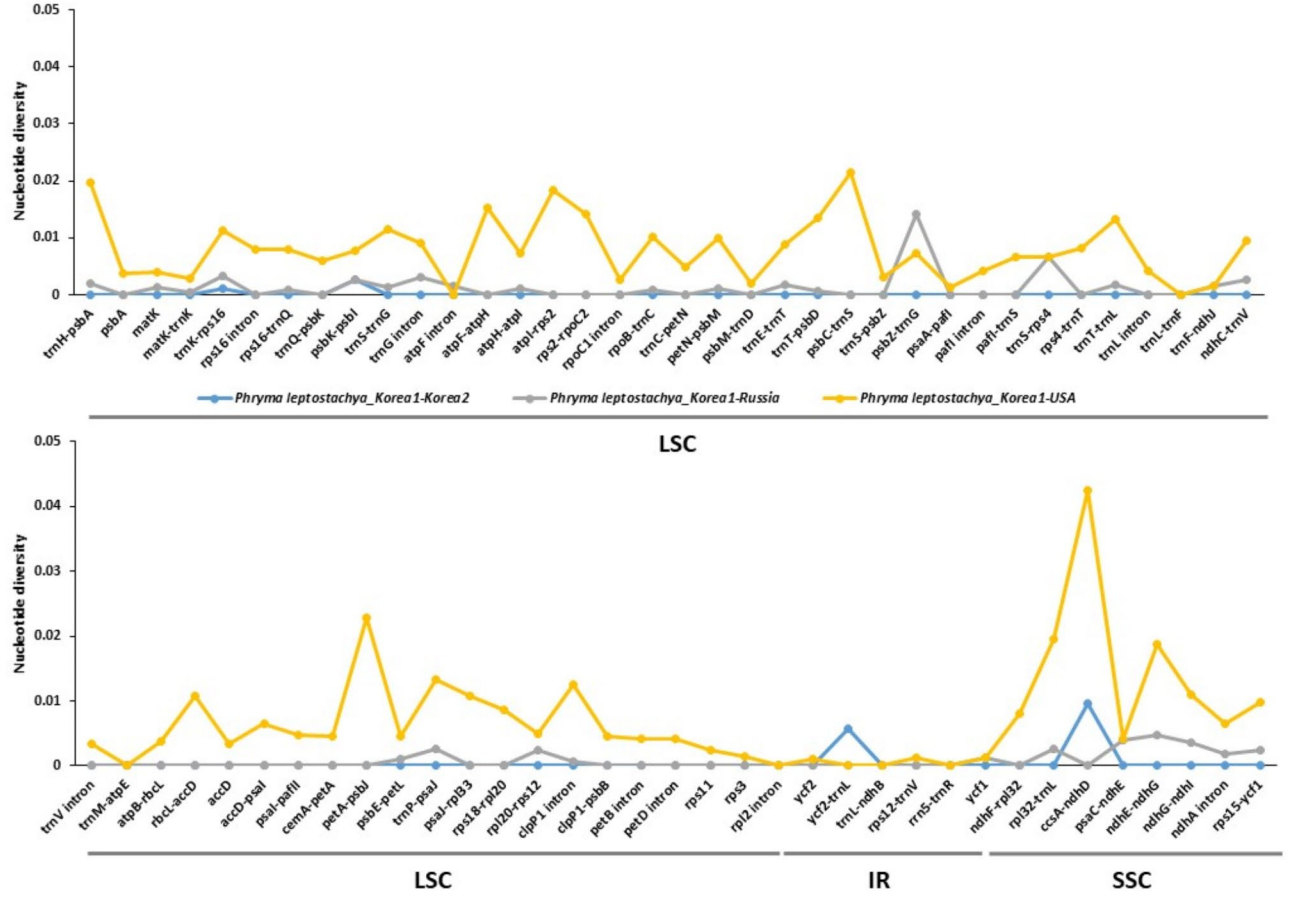

Fig. 4 Comparison of nucleotide diversity (Pi value) among the four plastomes. Each was compared to *Phryma leptostachya*_Korea1 as the reference: *P. leptostachya*_Korea2, *P. leptostachya*_Russia, and *P. leptostachya*_USA

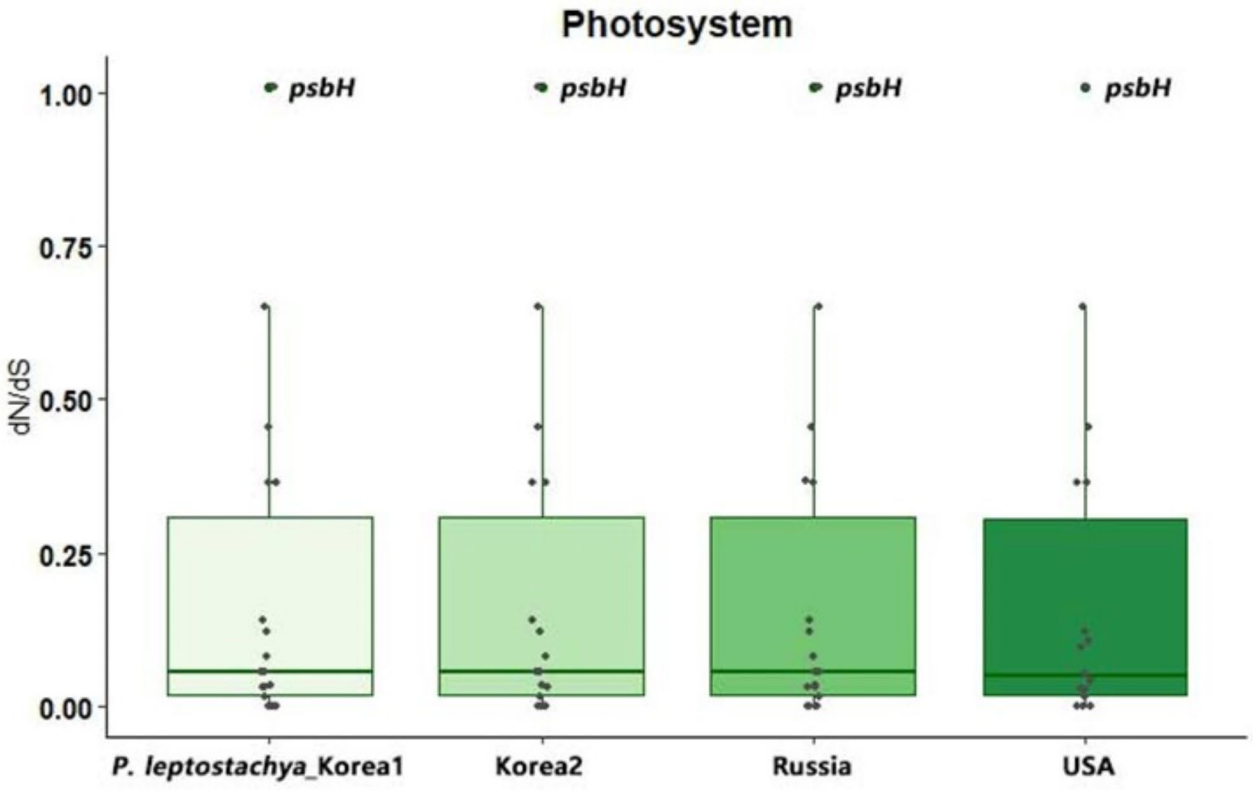

Fig. 5 Boxplot of dN/dS ratios for photosystem-related genes. Each dot represents the dN/dS value for an individual gene

Phylogenetic analysis

To confirm this phylogenetic relationship, we constructed a phylogenetic tree using ML and BI methods, including 15 Phrymaceae species, two Paulowniaceae species, and three Mazaceae species. In this study, *Lancea hirsuta*, *L. tibetica* and *Dodartia orientalis* were selected as outgroups (Fig. S8 and Table S3). The phylogenetic relationships observed in the results of both the ML and BI trees were identical, with robust support (ML = 100, PP = 1.0) for the majority of nodes. The tree was divided into three families, with members of the same genus clustered within each family. The branch nodes that divided *P. leptostachya*_Korea1, *P. leptostachya*_Russia, and *P. leptostachya*_USA were strongly supported (ML = 100, PP = 1.0). *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2 were very similar to each other. Considering *P. leptostachya*_Korea1 as the reference point, the distances were observed to be greater in the order of *P. leptostachya*_Russia > *P. leptostachya*_USA.

Divergence time estimation and ancestral state inference

We performed dating analysis using 78 PCGs from 20 plants, including *P. leptostachya* plastomes and outgroups. The results showed that the divergence between Phrymaceae and Paulowniaceae occurred at 57.64 MYA (95% HPD: 57.05–58.26), and *Phryma* diverged earlier within Phrymaceae than the two other genera (*Diplacus* and *Erythranthe*) at 47.77 MYA (95% HPD: 47.40–48.12) (Fig. 6). These divergent events occurred during the Paleocene and Eocene. Within *Phryma*, *P. leptostachya*_USA diverged first at 5.25 MYA (95% HPD: 4.89–5.59) during the Pliocene, followed by *P. leptostachya*_Russia at 0.87 MYA (95% HPD: 0.71–1.03) during the Quaternary period. The divergence of *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2 occurred more recently, at 0.11 MYA (95% HPD: 0.05–0.16), compared to other *P. leptostachya*.

Furthermore, using S-DIVA implemented in RASP (Reconstruct Ancestral State in Phylogenies), we inferred ancestral states and historical biogeographic patterns (Fig. 7). The distribution range was divided into three regions: (A) East Asia, (B) North America, and (C) the Russian Far East. Mazaceae was inferred to have originated in Region (A). The ancestral state of Paulowniaceae was inferred as AB (East Asia and North America), indicating its historical distribution between Regions (A) and (B), while the ancestral state of Phrymaceae was inferred as B (North America). Within Phrymaceae, the ancestral state of *P. leptostachya* was inferred as ABC (East Asia, North America, and the Russian Far East), suggesting its origin in Region (B), with subsequent dispersal and vicariance events. Our analysis identified three dispersal

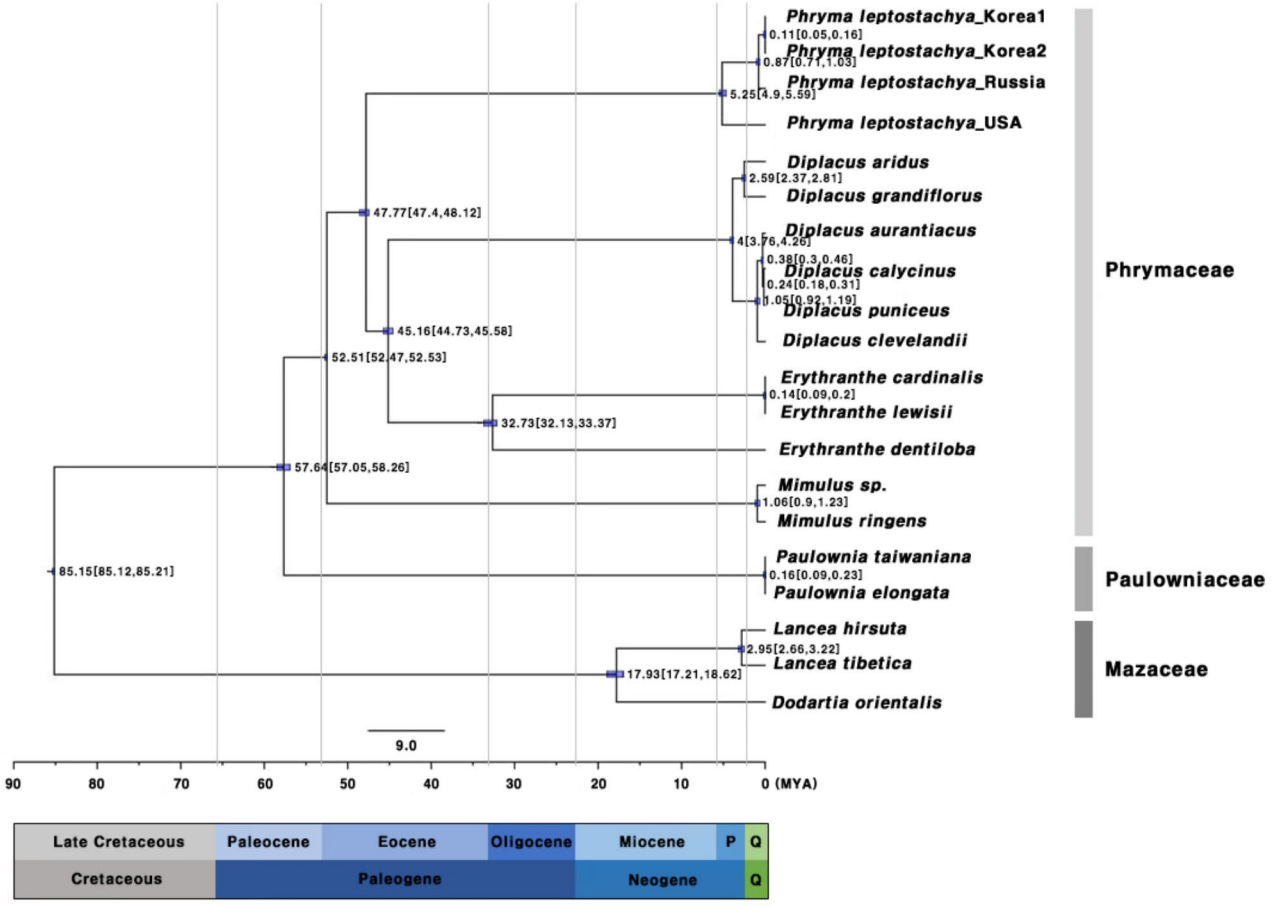

Fig. 6 Chronogram inferred from 78 protein-coding genes (PCGs). The molecular clock tree was constructed using the BEAST program. The numbers on the branches represent the mean age of divergence times (million years ago, MYA), with the 95% highest posterior density (HPD) confidence intervals for each node shown in the figure. Q: Quaternary includes the Holocene and Pleistocene, P: Pliocene

and three vicariance events. The first dispersal event involved migration from East Asia (A) to North America (B) as Paulowniaceae and Mazaceae diverged. The second dispersal event occurred during the divergence of Phrymaceae into three genera: *Diplacus*, *Erythranthe*, and *Phryma*. The third dispersal event occurred within *Phryma*, from North America (B) to East Asia (A) and the Russian Far East (C). Importantly, vicariance events occurred twice within *Phryma*: first from North America (B) to the Russian Far East (C), and then from the Russian Far East (C) to East Asia (A). These results strongly supported the hypothesis that *P. leptostachya* originated in North America and subsequently migrated to East Asia and the Russian Far East.

Haplotype network

To support the evidence of intercontinental differences, we constructed a haplotype network using *matK* sequences from 21 *P. leptostachya* samples (Fig. S9). This analysis identified three haplotypes: H1, H2, and H3. Compared to H1, which served as the reference, H2 had one substitution, while H3 had three substitutions. H1 was identical to the *matK* sequence found in *P. leptostachya* in Korea. H2 matched the *matK* sequence of *P. leptostachya*_Russia, and H3 was identical to the *matK* sequence of *P. leptostachya*_USA. Among the 21 samples, 11 were classified as H1, 5 as H2, and 5 as H3. H1 included samples distributed across East Asia, the Russian Far East, and North America. In contrast, H2 was found exclusively in East Asian samples, whereas H3 was observed only in North American samples.

Discussion

Features of *Phryma leptostachya* plastomes

We determined and analyzed the plastomes of *P. leptostachya*. The plastomes exhibited a typical quadripartite structure, similar to that of *P. leptostachya* var. *asiatica*, as previously reported, including the gene number and GC content [63]. The *ycf15* gene, often pseudogenized in many plants due to early stop codons [69, 70], was not identified. The RSCU value represents synonymous codon usage bias, indicating that the same amino acid is

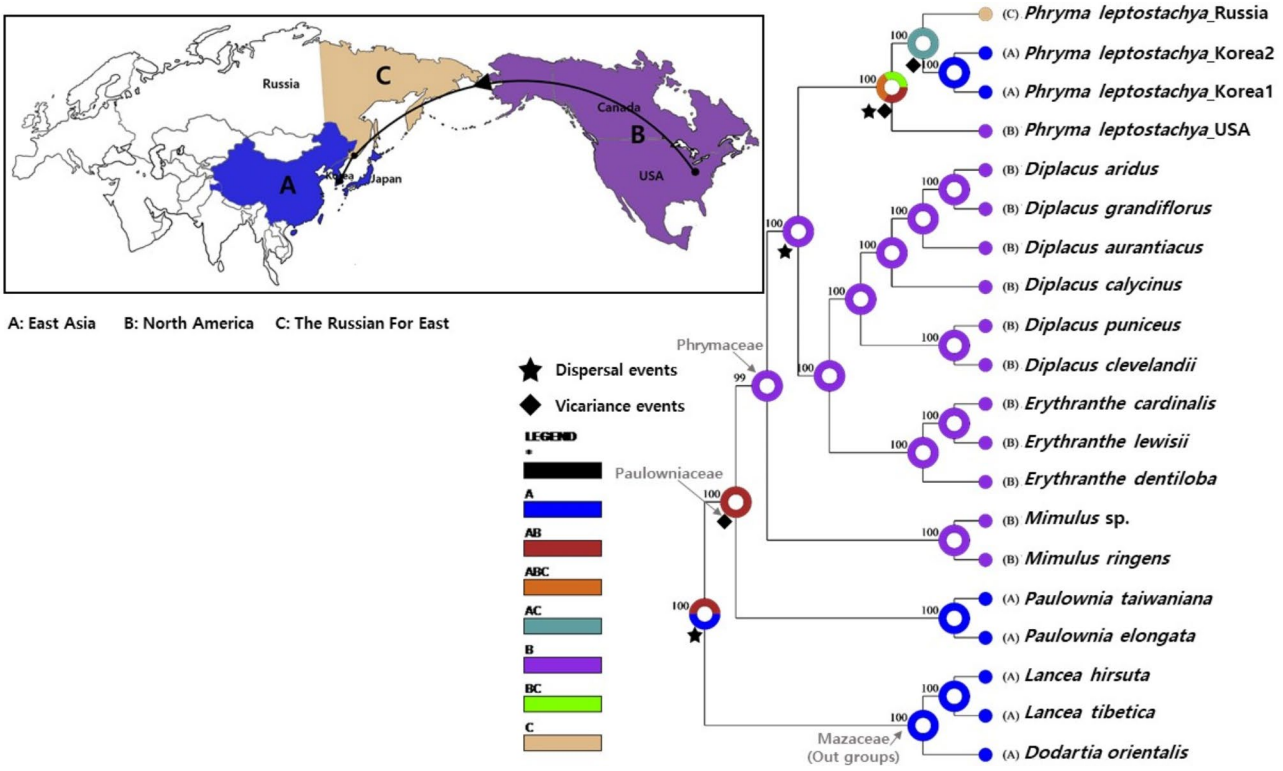

Fig. 7 Reconstruction of ancestral areas within the genus *Phryma* in Phrymaceae using the S-DIVA method. Star shapes represent dispersal events, whereas rhombus shapes represent vicariance events. Geographical regions are indicated by letters: A for East Asia, B for North America, and C for the Russian Far East. The upper left figure shows the sampling location of *P. leptostachya*, with black dots to show the path of movement

encoded by multiple codons. Leucine was the most commonly used codon, whereas cysteine was the least common (Fig. S4). When the RSCU value was greater than 1, most codons ended with A or T; when it was less than 1, the third position ended with G or C (Fig. S5). These findings have also been reported for many other plants [71, 72].

Correlation of repeat lengths with plastome size in *P. leptostachya*

SSRs are widely utilized for diversity studies and genetic and species identification, particularly in plants [73, 74]. In this study, the majority of SSRs identified were mononucleotide repeats within IGS regions, with a high abundance of A/T motifs (Fig. 2). This is likely due to the higher ratio of A/T in the plastome compared to that of G/C, a phenomenon reported in many plant groups [75–78]. Additionally, this finding is consistent with observations in other highly rearranged angiosperm genomes, where most repeats are located in the IGS region [79–83]. The repeat sequence lengths in *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2 were identical except for the SSRs. *P. leptostachya*_Russia exhibited shorter overall repeat lengths compared to *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2, whereas *P. leptostachya*_USA had the shortest repeat lengths. The plastome lengths followed a similar order: *P. leptostachya*_USA had the shortest plastome, followed by *P. leptostachya*_Russia and *P. leptostachya*_Korea. This pattern suggests that variations in repeat lengths are associated with differences in genome length.

Genomic variability revealed through the plastome

The IR is typically the most conserved region in the plastome compared to the LSC and SSC regions. The contraction and expansion of the IR in angiosperms are sometimes responsible for variations in genome size [84, 85]. We identified an IR region with a length of 25,373 bp in *P. leptostachya*_USA, which was similar in *P. leptostachya*_Korea1, *P. leptostachya*_Korea2, and *P. leptostachya*_Russia, ranging from 25,391 to 25,393 bp. The IRa/SSC and SSC/IRb junctions are highly conserved, whereas variability is observed at the LSC/IRa and IRb/LSC junctions. In particular, compared to other *P. leptostachya* variants, *P. leptostachya*_USA has an extended IR region, indicating structural changes (Fig. S6). The expansion and contraction of IR regions appear to play a role in determining genome length and are commonly observed in evolutionary processes [86, 87]. These changes may have affected the overall genome structure and evolutionary dynamics.

In this study, we compared the plastomes of *P. leptostachya* variants (Fig. 3). Variable regions were primarily observed in the LSC and SSC regions, whereas the IR region, which contains four highly conserved rRNA genes, showed fewer variations. Additionally, more variations were noted in the non-coding regions compared to the coding regions. Based on comparisons of *P. leptostachya*_Korea1 with *P. leptostachya*_Korea2, *P. leptostachya*_Russia, and *P. leptostachya*_USA, high nucleotide diversity was observed in the IGS regions (Fig. 4). The Pi values of *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2 were mostly zero, indicating similarity. In *P. leptostachya*_Russia, differences were observed in regions such as *psbZ-trnG* and *trnS-rps4*. Additionally, marked distinctions were noted in non-coding regions, such as *ccsA-ndhD*, *petA-psbJ*, and *psbC-trnS*, in *P. leptostachya*_USA. Notably, *P. leptostachya*_USA exhibited a higher level of diversity, which is unusual for a single species. This degree of variation suggests variation at the species level. Various regions, such as *ccsA-ndhD*, *petA-psbJ*, *psbC-trnS*, and *psbZ-trnG*, have been reported in other species, including *Atractylodes*, *Parrotia*, and *Hydrocotyle* [88–90]. Variations in non-coding regions, particularly in IGS regions, are valuable for distinguishing species and understanding genetic differences among them. These variable regions may contribute to the development of markers for species identification [91].

The dN/dS ratio is a valuable tool for assessing the evolutionary rate and selective pressure on specific PCGs. Most genes have undergone purifying selection due to their functional limitations, whereas others have undergone positive selection, leading to various environmental adaptations during evolution [92–96]. The *psbH* gene encodes a key Photosystem II subunit essential for photosynthesis. It is associated with CP47 and plays a crucial role in PSII assembly, stabilization, and electron transfer [97]. In shade-tolerant *Oryza* spp., positive selection of *psbH* can be interpreted as a genetic variation enhancing shade adaptation [98]. Additionally, in *Diospyros*, a dN/dS value greater than 1 for *psbH* in (sub) temperate species suggests that *psbH* plays a crucial role in adaptation to temperate regions [99]. A comparison of four *P. leptostachya* samples with *E. dentiloba* revealed positive selection exclusively in *psbH* ($\omega = 1.0089$) (Fig. 5, S7 and Table S11). Since the dN/dS value is slightly greater than 1, it suggests positive selection acting on *psbH*. Positive selection on *psbH* is estimated to have arisen from a divergence with a common ancestor approximately 42.77–52.51 MYA during the Eocene. Additionally, the identical dN/dS values suggest that adaptive changes in *psbH* have already been established in *P. leptostachya*. Environmental changes likely occur before differentiation, triggering the adaptive evolution of photosynthesis-related genes in response to these pressures [100].

Phylogenetic perspectives of *P. leptostachya*: distinct separation

To date, there have been few phylogenetic studies of *Phryma* using plastomes. In this study, we utilized 20 plastomes, including *P. leptostachya*. We constructed a phylogenetic tree using both ML and BI analyses, with three Mazaceae species included as outgroups (Fig. S8). Phrymaceae, Paulowniaceae, and Mazaceae have previously been classified phylogenetically based on plastome data [65]. Using ITS, ETS, and *trnL/F* sequence data to analyze phylogenetic relationships, we treated *P. leptostachya* from North America and *P. leptostachya* from East Asia as distinct varieties and confirmed considerable divergence between them [8]. Additionally, a phylogenetic tree constructed from a combination of ITS, *rps16*, and *trnL-F* sequences identified two major clades within *Phryma*: one consisting of populations from Eastern North America (ENA) and the other consisting of populations from Eastern Asia (EA) [15]. Consistent with previous research, our study confirmed a distinct phylogenetic separation between plastomes from North America and East Asia, despite their high morphological similarity. Specifically, *P. leptostachya*_Korea1 and *P. leptostachya*_Korea2 formed monophyletic groups, whereas *P. leptostachya*_Korea1 and *P. leptostachya*_Russia suggest the possibility of forming paraphyletic groups. Among the *P. leptostachya* variants examined, *P. leptostachya*_USA was the first to diverge phylogenetically. Notably, in this study, we also observed that *P. leptostachya* collected from the Far East region of Russia, was positioned between *P. leptostachya*_Korea and *P. leptostachya*_USA. Considering the divergence times reported in previous studies [15], our results suggest that the EA populations are more closely related to *P. leptostachya* from Korea and Russia, whereas the ENA populations are more closely related to *P. leptostachya* from the USA.

Migration of *P. leptostachya* via the Bering Land Bridge: a connector between East Asia and North America

We examined divergence times to support our findings on the phylogenetic relationships within *P. leptostachya* (Fig. 6). We found that the divergence between *P. leptostachya*_USA and the other *P. leptostachya* variants occurred approximately 5.25 million years ago (MYA), with subsequent divergences around 0.87 MYA and 0.11 MYA within the remaining samples. Previous studies estimated the divergence times of *P. leptostachya* var. *asiatica* from EA and *P. leptostachya* var. *leptostachya* in ENA to range from 3.68 ± 2.25 to 5.23 ± 1.37 MYA based on *rps16* and *trnL-F* sequence data [15]. Additionally, divergence estimates for the *Phryma* and its sister group range from 32.32 ± 4.46 to 49.35 ± 3.18 MYA [15], aligning with our estimate of divergence during the Eocene, approximately 47.77 MYA. Consequently, our results

are consistent, indicating that *P. leptostachya*_USA is closely related to ENA populations. Additionally, the EA populations were similar to *P. leptostachya*_Korea and *P. leptostachya*_Russia. However, *P. leptostachya*_Russia appears to have diverged from *P. leptostachya*_Korea approximately 0.87 MYA, indicating the presence of evolutionary differences exist between *P. leptostachya*_Korea and *P. leptostachya*_Russia.

The divergence of Paulowniaceae and Phrymaceae occurred during the Paleocene, marked by an initial vicariance event that led to their migration from East Asia (A) to North America (B). Within Phrymaceae, two dispersal and two vicariance events occurred; both vicariance events occurred within *Phryma*. One vicariance event occurred during the Pliocene, and the other occurred during the Quaternary, from Far East Russia to East Asia (Fig. 7). The haplotype analysis of *P. leptostachya* confirmed a discontinuous distribution driven by historical vicariance and dispersal events (Fig. S9). Specifically, haplotype H1 was found in East Asia, the Russian Far East, and North America; H2 was restricted to East Asia; and H3 was confined to North America. The results showed that H1, H2, and H3 corresponded to *P. leptostachya* from Korea, Russia, and the USA, respectively, highlighting similar patterns of discontinuous distribution. The observed distribution patterns were likely influenced by historical climate changes that affected global temperatures, precipitation patterns, and sea levels. Vicariance events in the Pliocene occurred under warmer and wetter conditions, which facilitated the movement of plants between regions [101]. The Pleistocene experienced significant fluctuations in plant communities owing to alternating glaciations and interglacial periods, which profoundly affected their distribution [102]. The vicariance event that separated *P. leptostachya*_Russia from *P. leptostachya*_Korea likely resulted from geographical or environmental changes and potentially led to independent evolutionary trajectories over time. Our S-DIVA results support a dispersal scenario from North America to East Asia that was likely facilitated by the Bering Land Bridge. Furthermore, we believe that *P. leptostachya* in Russia likely played a bridging role between North America and East Asia. This bridge intermittently connected the two continents during periods of lower sea level between the Miocene and Pliocene, playing a crucial role in enabling the intercontinental migration of various species, including plants [103, 104].

Additionally, the separation patterns between East Asia and North America are generally associated with the use of the Bering Land Bridge and the North Atlantic Land Bridge [26, 105]. The North Atlantic Land Bridge, which connects Europe to North America, existed from the Late Paleocene to the Early Eocene. In contrast, the Bering Land Bridge is believed to have existed from the Palaeocene to the Miocene and temporarily reopened throughout the Pleistocene [106]. The Bering Land Bridge may have been more significant than the North Atlantic Land Bridge for temperate taxa [26], thus reinforcing our hypothesis of migration via the Bering Land Bridge. Moreover, *P. leptostachya* relies on fruits rather than seeds for animal-mediated dispersal, as their fruits have characteristics that allow them to adhere well to animals [107]. This dispersal method can aid in overcoming geographical barriers and facilitating long-distance movement, thus complementing the role of land bridges in enabling plants to spread across continents. Similar migration patterns were observed in other species. For instance, *Circaea* migrated from Eurasia to North America via the Bering Land Bridge, as supported by analyses of partial sequences (*petB-petD*, *rpl16*, and *trnL-F*) and nrITS [108]. *Pseudotsuga* originated in North America and subsequently migrated to East Asia during the Early Miocene [109]. Additionally, *Chamaecyparis* experienced geographical separation between North America and East Asia on at least two occasions [110]. Similar migration patterns have been observed in *Penthorum* [111], *Cornus* [112], *Sassafras* [16], and *Saxifraga rivularis* [113].

Our analysis estimated the divergence of *P. leptostachya* from Korea, Russia, and the USA, suggesting that *P. leptostachya* likely migrated from the USA to Korea though the Bering Land Bridge. This migration route likely played a crucial role in shaping the observed genetic differentiation. During this migration, *P. leptostachya* underwent prolonged geographical isolation and adaptation to distinct environmental conditions, which further contributed to its genetic divergence. These genetic differences can be attributed to geographical barriers and environmental factors that restricted gene flow and fostered distinct evolutionary pathways.

Conclusion

In this study, we examined the plastomes of *P. leptostachya* variants from Korea, Russia, and the USA. Sequence differences were observed among *P. leptostachya* variants, particularly in the IGS regions, with *P. leptostachy*a_USA showing IR expansion. We constructed phylogenetic relationships that revealed clear separation within *P. leptostachya* and estimated divergence times approximately 5.25, 0.87, and 0.11 MYA for *P. leptostachya_Korea*, *P. leptostachya_Russia*, and *P. leptostachya_USA*, respectively, yielding high-resolution results. Additionally, our study provides insights into the evolutionary dynamics of Phrymaceae, revealing key dispersal, vicariance events, and intercontinental migration patterns. Notably, the biogeographic results suggest that *P. leptostachya* migrated from North America to East Asia, with *P. leptostachya*_Russia likely serving

as a bridge during this process. Migration, dispersal, and vicariance events, along with geographical isolation, led to genomic variation that may have contributed to divergence into separate species. Therefore, we propose that *P. leptostachya* should be recognized as a separate species rather than a single species. This study highlights the geographic differences among *P. leptostachya* plastomes and enhances our understanding of the evolutionary trajectory of *Phryma*.

Abbreviations

bp	Base pairs
BR	Meise Botanic Garden Herbarium
BWA	Burrows-Wheeler Aligner
CDS	Coding Sequences
DnaSP	DNA Sequence Polymorphism
GC	Guanine-Cytosine
IR	Inverted Repeat
IRa	Inverted Repeat A
IRb	Inverted Repeat B
KB	National Institute of Biological Resources
L	Naturalis Biodiversity Center
LSC	Large Single Copy
MISA	Microsatellite Identification Tool
mVISTA	A program used to visualize sequence and structural differences in genomic data
MYA	Million Years Ago
OGDRAW	Online Gene Data Representation and Annotation Web Tool
P. leptostachya	*Phryma leptostachya*
RSCU	Relative Synonymous Codon Usage
SAM	Sequence Alignment/Map
SSRs	Simple Sequence Repeats

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06272-9.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We express our gratitude to the directors of the herbaria BR, KB, and L for granting us permission to examine the specimens. We would also like to extend our special thanks to Dr. J.D. Jung from the Northeastern Asia Biodiversity Institute and Prof. Dr. P.D. Cantino from Ohio University for generously donating valuable samples for this study.

Author contributions

Yeseul Kim: Writing– original draft, Software, Methodology, Formal analysis, Data curation; Sumin Jeong: Writing– original draft, Data curation, Visualization; Inkyu Park: Writing– review & editing, Conceptualization, Supervision; Hye-Kyoung Moon: Writing - review & editing– Conceptualization, Resources, Funding acquisition, Project administration, Supervision.

Funding

This study was supported by National Research Foundation of South Korea (NRF-2014R1A1A2058195). And this research was supported by Global - Learning & Academic research institution for Master's·PhD students, and Postdocs (LAMP) Program of the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. RS2024-00444460).

Data availability

The newly sequenced chloroplast genome sequences in this study were deposited in the NCBI GenBank database under the accession numbers (PQ443460, PQ443461, and PQ443462).

Declarations

Ethics approval and consent to participate

Fresh samples of *Phryma leptostachya* were directly collected from Gayasan Mt., Hapcheon-gun, Gyeongsangnam-do, South Korea (KHUM20140008: 35.800323°N, 128.098229°E) by H.K. Moon. Additionally, samples from Vladivostok, Primorsky Krai, Russia (KHUJ20140001: 43.177034°N, 131.946176°E) and Athens County, Ohio, USA (KHUC20141453: 39.344397°N, 82.087248°W) were kindly provided by collaborators. These voucher specimens were deposited in the herbarium of Kyung Hee University. For barcoding analysis, dried samples were obtained from herbarium specimens at the Meise Botanic Garden in Belgium (BR), National Institute of Biological Resources in South Korea (KB), and Naturalis Biodiversity Center in the Netherlands (L), with appropriate permissions. This study adheres to all applicable institutional, national, and international guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Biology and Chemistry, Changwon National University, Changwon, Republic of Korea
[2]Department of Biology, Kyung Hee University, Seoul, Republic of Korea

Received: 30 October 2024 / Accepted: 17 February 2025
Published online: 04 March 2025

References

1. Li H. Floristic relationships between Eastern Asia and Eastern America. Trans Ames Philos Soc. 1952;42:371–429.
2. Hong D, Phryma JW. In: Cui H, Hong D, editors. Flora of China. Beijing: Science; St. Louis: Missouri Botanical Garden Press;, 2011. pp. 493–4.
3. Raven PH, Axelrod DI. Angiosperm biogeography and past continental movements. Ann Mo Bot Gard. 1974;61:539–673.
4. Dahlgren RMT. A revised system of classification of the angiosperms. Bot J Linn Soc. 1980;80:91–124.
5. Lu AM. A preliminary cladistic study of the families of the superorder lamiiflorae. Bot J Linn Soc. 1990;103:36–57.
6. Chadwell TB, Wagstaff SJ, Cantino PD. Pollen morphology of *Phryma* and some putative relatives. Syst Bot. 1992;17:210–9.
7. Oxelman B, Kornhall P, Olmstead RG, Bremer B. Further disintegration of Scrophulariaceae. Taxon. 2005;54:411–25.
8. Beardsley PM, Olmstead RG. Redefining phrymaceae: the placement of *Mimulus*, tribe mimuleae, and *Phryma*. Am J Bot. 2002;89:1093–102.
9. Liu B, Tan YH, Liu S, Olmstead RG, Min DZ, Chen ZD, et al. Phylogenetic relationships of *Cyrtandromoea* and *Wightia* revisited: A new tribe in Phrymaceae and a new family in Lamiales. J Syst Evol. 2020;58:1–17.
10. Baker WR, Nesom GL, Beardsley PM, Fraga NS. A taxonomic conspectus of phrymaceae: A narrowed circumscription for *Mimulus*, new and resurrected genera, and new names and combinations. Phytoneuron. 2012;39:1–60.
11. Parks CR, Wendel JF. Molecular divergence between Asian and North American species of *Liriodendron* (Magnoliaceae) with implications for interpretation of fossil floras. Am J Bot. 1990;77:1243–56.
12. Hoey MT, Parks CR. Isozyme divergence between Eastern Asian, North American, and Turkish species of *Liquidambar* (Hamamelidaceae). Am J Bot. 1991;78:938–47.
13. Wen J, Jansen RK. Morphological and molecular comparisons of *Campsis grandiflora* and *C. radicans* (Bignoniaceae), an Eastern Asian and Eastern North American vicariad species pair. Pl Syst Evol. 1995;196:173–83.

14. Wen J, Jansen RK, Kilgore K. Evolution of the Eastern Asian and Eastern North American disjunct genus *Symplocarpus* (Araceae): insights from Chloroplast DNA restriction site data. Biochem Syst Ecol. 1996;24:735–47.
15. Nie ZL, Sun H, Beardsley PM, Olmstead RG, Wen J. Evolution of biogeographic disjunction between Eastern Asia and Eastern North America in *Phryma* (Phrymaceae). Am J Bot. 2006;93:1343–56.
16. Nie ZL, Wen J, Sun H. Phylogeny and biogeography of *Sassafras* (Lauraceae) disjunct between Eastern Asia and Eastern North America. Plant Syst Evol. 2007;267:191–203.
17. Guo ZT, Sun B, Zhang ZS, Peng SZ, Xiao GQ, Ge JY, et al. A major reorganization of Asian climate regime by the early miocene. Clim Past Discuss. 2008;4:535–84.
18. Jiang Y, Gao M, Meng Y, Wen J, Ge XJ, Nie ZL. The importance of the North Atlantic land bridges and Eastern Asia in the post-Boreotropical biogeography of the Northern hemisphere as revealed from the poison Ivy genus (*Toxicodendron*, Anacardiaceae). Mol Phylogenet Evol. 2019;139:106561.
19. Lee NS. Allozyme divergence in *Phryma leptostachya* (Phrymaceae). Korean J Pl Taxon. 1990;20:147–56.
20. Lee NS, Sang T, Crawford DJ, Yeau SH, Kim SC. Molecular divergence between disjunct taxa in Eastern Asia and Eastern North America. Am J Bot. 1996;83:1373–8.
21. Dong W, Xu C, Liu Y, Shi J, Li W, Suo Z. Chloroplast phylogenomics and divergence times of *Lagerstroemia* (Lythraceae). BMC Genomics. 2021;22:434.
22. Shang C, Li E, Yu Z, Lian M, Chen Z, Liu K, et al. Chloroplast genomic resources and genetic divergence of endangered species *Bretschneidera sinensis* (Bretschneideraceae). Front Ecol Evol. 2022;10:873100.
23. Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. Genomics. 2016;107:1–8.
24. Ji Y, Yang L, Chase MW, Liu C, Yang Z, Yang J, et al. Plastome phylogenomics, biogeography, and clade diversification of *Paris* (Melanthiaceae). BMC Plant Biol. 2019;19:543.
25. Shi W, Song W, Chen Z, Cai H, Gong Q, Liu J, et al. Comparative Chloroplast genome analyses of diverse *Phoebe* (Lauraceae) species endemic to China provide insight into their phylogeographical origin. PeerJ. 2023;11:e14573.
26. Tiffney BH. The eocene North Atlantic land Bridge: its importance in tertiary and modern phytogeography of the Northern hemisphere. J Arnold Arbor. 1985;66:243–73.
27. Jansen RK, Raubeson LA, Boore JL, dePamphilis CW, Chumley TW, Haberle RC, et al. Methods for obtaining and analyzing whole Chloroplast genome sequences. Methods Enzymol. 2005;395:348–84.
28. Moore MJ, Bell CD, Soltis PS, Soltis DE. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc Natl Acad Sci U S A. 2007;104:19363–8.
29. Yi DK, Kim KJ. Complete Chloroplast genome sequences of important oilseed crop *Sesamum indicum* L. PLoS ONE. 2012;7:e35872.
30. Givnish TJ, Spalink D, Ames M, Lyon SP, Hunter SJ, Zuluaga A, et al. Orchid historical biogeography, diversification, Antarctica and the paradox of Orchid dispersal. J Biogeogr. 2016;43:1905–16.
31. Cheng Y, Zhang L, Qi J, Zhang L. Complete chloroplast genome sequence of *Hibiscus cannabinus* and comparative analysis of the Malvaceae family. Front Genet. 2020;11:227.
32. Park I, Yang S, Song JH, Moon BC. Dissection for floral micromorphology and plastid genome of valuable medicinal borages *Arnebia* and *Lithospermum* (Boraginaceae). Front Plant Sci. 2020;11:606463.
33. Park I, Choi B, Weiss-Schneeweiss H, So S, Myeong HH, Jang TS. Comparative analyses of complete Chloroplast genomes and karyotypes of allotetraploid *Iris Koreana* and its putative diploid parental species (*Iris* series *Chinenses*, Iridaceae). Int J Mol Sci. 2022;23:10929.
34. Liang H, Zhang Y, Deng J, Gao G, Ding C, Zhang L, et al. The complete Chloroplast genome sequences of 14 *Curcuma* species: insights into genome evolution and phylogenetic relationships within Zingiberales. Front Genet. 2020;11:802.
35. Wang J, Qian J, Jiang Y, Chen X, Zheng B, Chen S, et al. Comparative analysis of Chloroplast genome and new insights into phylogenetic relationships of *Polygonatum* and tribe polygonateae. Front Plant Sci. 2022;13:882189.
36. Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc. 2006;1:2320–5.
37. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014;30:2114–20.
38. Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, et al. GetOrganelle: a fast and versatile toolkit for accurate de Novo assembly of organelle genomes. Genome Biol. 2020;21:241.
39. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq– versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45:W6–11.
40. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9.
41. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv. 2013;13:13033997.
42. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and samtools. Bioinformatics. 2009;25:2078–9.
43. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47:W59–64.
44. Tamura K, Stecher G, Mega KS. Molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38:3022–7.
45. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, et al. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 2016;44:W147–53.
46. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 2004;32:W273–9.
47. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34:3299–302.
48. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33:2583–5.
49. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29:4633–42.
50. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–80.
51. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24:1586–91.
52. Group TAP, Chase MW, Christenhusz MJM, Fay MF, Byng JW, Judd WS, et al. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc. 2016;181:1–20.
53. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. Nucleic Acids Res. 2002;30:3059–66.
54. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540–52.
55. Darriba D, Taboada GL, Doallo R, Posada D. JModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9:772.
56. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61:539–42.
57. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, et al. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. PLOS Comput Biol. 2019;15:e1006650.
58. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using tracer 1.7. Syst Biol. 2018;67:901–4.
59. Rambaut A. FigTree v1.4.2: tree drawing tool in computer program and documentation distributed by the author; 2014. http://tree.bio.ed.ac.uk/software/figtree
60. Chi X, Wang J, Gao Q, Zhang F, Chen S. The complete Chloroplast genomes of two *Lancea* species with comparative analysis. Molecules. 2018;23:602.
61. Zhao H, Li R, Shang F. The complete Chloroplast genome of *Paulownia elongata* and phylogenetic implications in Lamiales. Mitochondrial DNA B. 2019;4:2067–8.
62. Li P, Lou G, Cai X, Zhang B, Cheng Y, Wang H. Comparison of the complete plastomes and the phylogenetic analysis of *Paulownia* species. Sci Rep. 2020;10:2225.
63. Jo S, Kim HW, Kim YK, Cheon SH, Joo MJ, Hong JR, et al. Three complete plastome sequences from the families of lamiaceae, Mazaceae, and Phrymaceae (Lamiales). Mitochondrial DNA B Resour. 2021;6:224–6.
64. Zeng S, Li J, Yang Q, Wu Y, Yu J, Pei X, et al. Comparative plastid genomics of Mazaceae: focusing on a new recognized genus. Puchiumazus Planta. 2021;254:99.

65. Zhao F, Chen YP, Salmaki Y, Drew BT, Wilson TC, Scheen AC, et al. An updated tribal classification of Lamiaceae based on plastome phylogenomics. BMC Biol. 2021;19:2.
66. Yu Y, Harris AJ, Blair C, He X. RASP. RASP (Reconstruct ancestral state in Phylogenies): a tool for historical biogeography. Mol Phylogenet Evol. 2015;87:46–9.
67. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 2007;35:W71–4.
68. Leigh JW, Bryant D. Popart: full-feature software for haplotype network construction. Methods Ecol Evol. 2015;6:1110–6.
69. Steane DA. Complete nucleotide sequence of the chloroplast genome from the Tasmanian blue gum, *Eucalyptus globulus* (Myrtaceae). DNA Res. 2005;12:215–20.
70. Li Y, Zhou JG, Chen XL, Cui YX, Xu ZC, Li YH, et al. Gene losses and partial deletion of small single-copy regions of the chloroplast genomes of two hemiparasitic *Taxillus* species. Sci Rep. 2017;7:12834.
71. Morton BR. Selection on the codon bias of chloroplast and cyanelle genes indifferent plant and algal lineages. J Mol Evol. 1998;46:449–59.
72. Wang L, Wuyun T, Du H, Wang D, Cao D. Complete chloroplast genome sequences of *Eucommia ulmoides*: genome structure and evolution. Tree Genet Genomes. 2016;12:1–15.
73. Saha MC, Cooper JD, Mian MAR, Chekhovskiy K, May GD. Tall fescue genomic SSR markers: development and transferability across multiple grass species. Theor Appl Genet. 2006;113:1449–58.
74. Jena KK, Mackill DJ. Molecular markers and their use in marker-assisted selection in rice. Crop Sci. 2008;48:1266–76.
75. Melotto-Passarin DM, Tambarussi EV, Dressano K, De Martin VF, Carrer H. Characterization of chloroplast DNA microsatellites from *Saccharum* spp and related species. Genet Mol Res. 2011;10:2024–33.
76. Martin G, Baurens FC, Cardi C, Aury JM, D'Hont A. The complete chloroplast genome of banana (*Musa acuminata*, Zingiberales): insight into plastid monocotyledon evolution. PLoS ONE. 2013;8:e67350.
77. Curci PL, De Paola D, Danzi D, Vendramin GG, Sonnante G. Complete chloroplast genome of the multifunctional crop globe artichoke and comparison with other Asteraceae. PLoS ONE. 2015;10:e0120589.
78. Asaf S, Khan AL, Khan AR, Waqas M, Kang SM, Khan MA, et al. Complete chloroplast genome of *Nicotiana otophora* and its comparison with related species. Front Plant Sci. 2016;7:843.
79. Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, et al. The complete chloroplast genome sequence of *Pelargonium* x *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol. 2006;23:2175–90.
80. Lee HL, Jansen RK, Chumley TW, Kim KJ. Gene relocations within chloroplast genomes of *Jasminum* and *Menodora* (Oleaceae) are due to multiple, overlapping inversions. Mol Biol Evol. 2007;24:1161–80.
81. Yue F, Cui L, dePamphilis CW, Moret BM, Tang J. Gene rearrangement analysis and ancestral order inference from chloroplast genomes with inverted repeat. BMC Genomics. 2007;9(Suppl 1):1–9.
82. de Cambiaire JC, Otis C, Turmel M, Lemieux C. The chloroplast genome sequence of the green alga *Leptosira terrestris*: multiple losses of the inverted repeat and extensive genome rearrangements within the Trebouxiophyceae. BMC Genomics. 2007;8:213.
83. Haberle RC, Fourcade HM, Boore JL, Jansen RK. Extensive rearrangements in the chloroplast genome of *Trachelium caeruleum* are associated with repeats and tRNA genes. J Mol Evol. 2008;66:350–61.
84. Plunkett GM, Downie SR. Expansion and contraction of the chloroplast inverted repeat in Apiaceae subfamily Apioideae. Syst Bot. 2000;25:648–67.
85. Raubeson LA, Peery R, Chumley TW, Dziubek C, Fourcade HM, Boore JL, et al. Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. BMC Genomics. 2007;8:174.
86. Song W, Chen Z, Shi W, Han W, Feng Q, Shi C, et al. Comparative analysis of complete chloroplast genomes of nine species of *Litsea* (Lauraceae): hypervariable regions, positive selection, and phylogenetic relationships. Genes (Basel). 2022;13:1550.
87. Cao J, Wang H, Cao Y, Kan S, Li J, Liu Y. Extreme reconfiguration of plastid genomes in Papaveraceae: rearrangements, gene loss, pseudogenization, IR expansion, and repeats. Int J Mol Sci. 2024;25:2278.
88. Zhang YY, Shi E, Yang ZP, Geng QF, Qiu YX, Wang ZS. Development and application of genomic resources in an endangered Palaeoendemic Tree, *Parrotia subaequalis* (Hamamelidaceae) From Eastern China. Front Plant Sci. 2018;9:246.
89. Xu W, Lu R, Li J, Xia M, Chen G, Li P. Comparative plastome analyses and evolutionary relationships of all species and cultivars within the medicinal plant genus *Atractylodes*. Ind Crops Prod. 2023;201:116974.
90. Yan R, Gu L, Qu L, Wang X, Hu G. New insights into phylogenetic relationship of *Hydrocotyle* (Araliaceae) based on plastid genomes. Int J Mol Sci. 2023;24:16629.
91. Zhao Y, Yin J, Guo H, Zhang Y, Xiao W, Sun C, et al. The complete chloroplast genome provides insight into the evolution and polymorphism of *Panax ginseng*. Front Plant Sci. 2014;5:696.
92. Sloan DB, Triant DA, Forrester NJ, Bergner LM, Wu M, Taylor DR. A recurring syndrome of accelerated plastid genome evolution in the angiosperm tribe *Sileneae* (Caryophyllaceae). Mol Phylogenet Evol. 2014;72:82–9.
93. Wang Y, Zhan DF, Jia X, Mei WL, Dai HF, Chen XT et al. Complete chloroplast genome sequence of *Aquilaria sinensis* (Lour.) Gilg and evolution analysis within the Malvales order. Front Plant Sci. 2016;7:280.
94. Fan WB, Wu Y, Yang J, Shahzad K, Li ZH. Comparative chloroplast genomics of Dipsacales species: insights into sequence variation, adaptive evolution, and phylogenetic relationships. Front Plant Sci. 2018;9:689.
95. Huang S, Ge X, Cano A, Salazar BGM, Deng Y. Comparative analysis of chloroplast genomes for five *Dicliptera* species (Acanthaceae): molecular structure, phylogenetic relationships, and adaptive evolution. PeerJ. 2020;8:e8450.
96. Yang J, Takayama K, Youn JS, Pak JH, Kim SC. Plastome characterization and phylogenomics of East Asian beeches with a special emphasis on *Fagus multinervis* on Ulleung Island. Korea Genes. 2020;11:1338.
97. Bock R. Structure, function, and inheritance of plastid genomes. In: Bock R, editor. Cell and molecular biology of plastids. Berlin: Springer; 2007. pp. 29–63.
98. Gao LZ, Liu YL, Zhang D, Li W, Gao J, Liu Y, et al. Evolution of *Oryza* chloroplast genomes promoted adaptation to diverse ecological habitats. Commun Biol. 2019;2:278.
99. Huang Y, Ma Q, Sun J, Zhou LN, Lai CJ, Li P, et al. Comparative analysis of *Diospyros* (Ebenaceae) plastomes: insights into genomic features, mutational hotspots, and adaptive evolution. Ecol Evol. 2023;13:e10301.
100. Wang Y, Wen F, Hong X, Li Z, Mi Y, Zhao B. Comparative chloroplast genome analyses of *Paraboea* (Gesneriaceae): insights into adaptive evolution and phylogenetic analysis. Front Plant Sci. 2022;13:1019831.
101. Salzmann U, Williams M, Haywood AM, Johnson ALA, Kender S, Zalasiewicz J. Climate and environment of a Pliocene warm world. Palaeogeogr Palaeoclimatol Palaeoecol. 2011;309:1–8.
102. Hampe A, Jump AS. Climate relicts: past, present, future. Annu Rev Ecol Evol Syst. 2011;42:313–33.
103. Wen J. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annu Rev Ecol Syst. 1999;30:421–55.
104. Wen J, Ickert-Bond SM. Evolution of the Madrean–Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. J Syst Evol. 2009;47:331–48.
105. Wen J, Ickert-Bond S, Nie ZL, Li R. Timing and modes of evolution of eastern Asian-North American biogeographic disjunctions in seed plants. In: Darwin's heritage today. proceedings of the Darwin 200 Beijing international conference. Long M, Gu H, Zhou Z, editors. Beijing: Higher Education Press; 2010. pp. 252–69.
106. Gladenkov AY, Oleinik AE, Marincovich L Jr, Barinov KB. A refined age for the earliest opening of Bering Strait. Palaeogeography Palaeoclimatology Palaeoecology. 2002;183:321–8.
107. Holm T. *Phryma leptostachya* L., A morphological study. Bot Gaz. 1913;56:306–18.
108. Xie L, Wagner WL, Ree RH, Berry PE, Wen J. Molecular phylogeny, divergence time estimates, and historical biogeography of *Circaea* (Onagraceae) in the Northern Hemisphere. Mol Phylogenet Evol. 2009;53:995–1009.
109. Wei XX, Yang ZY, Li Y, Wang XQ. Molecular phylogeny and biogeography of *Pseudotsuga* (Pinaceae): insights into the floristic relationship between Taiwan and its adjacent areas. Mol Phylogenet Evol. 2010;55:776–85.
110. Liao PC, Lin TP, Hwang SY. Reexamination of the pattern of geographical disjunction of *Chamaecyparis* (Cupressaceae) in North America and East Asia. Bot Stud. 2010;51:511–20.
111. Xiang QY, Soltis DE, Soltis PS, Manchester SR, Crawford DJ. Timing the eastern Asian–eastern North American floristic disjunction: molecular clock corroborates paleontological estimates. Mol Phylogenet Evol. 2000;15:462–72.
112. Xiang QY, Thomas DT, Zhang W, Manchester SR, Murrell Z. Species level phylogeny of the genus *Cornus* (Cornaceae) based on molecular and morphological evidence—implications for taxonomy and Tertiary intercontinental migration. Taxon. 2006;55:9–30.

113. Westergaard KB, Jørgensen MH, Gabrielsen TM, Alsos IG, Brochmann C. The extreme Beringian/Atlantic disjunction in *Saxifraga rivularis* (Saxifragaceae) has formed at least twice. J Biogeogr. 2010;37:1262–76.

Publisher's note