

# Plastome phylogenomics, biogeography, and clade diversification of Paris (Melanthiaceae)

Yunheng Ji (✉ [jiyh@mail.kib.ac.cn](mailto:jiyh@mail.kib.ac.cn))

Kunming Institute of Botany Chinese Academy of Sciences <https://orcid.org/0000-0002-1116-7485>

Lifang Yang

Kunming Institute of Botany Chinese Academy of Sciences

Mark W. Chase

Royal Botanic Gardens Kew

Changkun Liu

Kunming Institute of Botany Chinese Academy of Sciences

Zhenyan Yang

Kunming Institute of Botany Chinese Academy of Sciences

Jin Yang

Kunming Institute of Botany Chinese Academy of Sciences

Jun-Bo Yang

Kunming Institute of Botany Chinese Academy of Sciences

Ting-Shuang Yi

Kunming Institute of Botany Chinese Academy of Sciences

---

## Research article

**Keywords:** Plastid phylogenomics, biogeography, radiative diversification, cytonuclear discordance, large genome size, Parideae, Paris, Melanthiaceae, Trilliaceae

**Posted Date:** December 2nd, 2019

**DOI:** <https://doi.org/10.21203/rs.2.12669/v4>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on December 5th, 2019. See the published version at <https://doi.org/10.1186/s12870-019-2147-6>.

# Abstract

Background: *Paris* (Melanthiaceae) is an economically important but taxonomically difficult genus, which is unique in angiosperms because some species have extremely large nuclear genomes. Phylogenetic relationships within *Paris* have long been controversial. Based on complete plastomes and nuclear ribosomal DNA (nrDNA) sequences, this study aims to reconstruct a robust phylogenetic tree and explore historical biogeography and clade diversification in the genus. Results: All 29 species currently recognized in *Paris* were sampled. Whole plastomes and nrDNA sequences were generated by the genome skimming approach. Phylogenetic relationships were reconstructed using the maximum likelihood and Bayesian inference methods. Based on the phylogenetic framework and molecular dating, biogeographic scenarios and historical diversification of *Paris* were explored. Significant conflicts between plastid and nuclear datasets were identified, and the plastome tree is highly congruent with past interpretations of the morphology. Ancestral area reconstruction indicated that *Paris* may have originated in northeastern Asia and northern China, and has experienced multiple dispersal and vicariance events during its diversification. The rate of clade diversification has sharply accelerated since the Miocene/Pliocene boundary. Conclusions: Our results provide important insights for clarifying some of the long-standing taxonomic debates in *Paris*. Cytonuclear discordance may have been caused by ancient and recent hybridizations in the genus. The climatic and geological changes since the late Miocene, such as the intensification of Asian monsoon and the rapid uplift of Qinghai-Tibet Plateau, as well as the climatic fluctuations during the Pleistocene, played essential roles in driving range expansion and radiative diversification in *Paris*. Our findings challenge the theoretical prediction that large genome sizes may limit speciation.

## Background

*Paris* is a small genus that was once placed in Trilliaceae [1], but now in Melanthiaceae [2–6]. The genus comprises of *ca.* 29 species of understory perennial herbs that are continuously distributed across Eurasia [7–9]. With most species (24/29) occurring in China and Himalayas, *Paris* may have experienced significant species diversification in subtropical East Asia (21°–34° N) [7, 9, 10]. Most species of this genus are much-valued traditional medicinal herbs in China and neighboring countries due to their various therapeutic properties [11–13]. Among them, the rhizomes of *Paris polyphylla* var. *chinensis* and *P. polyphylla* var. *yunnanensis* (*Rhizome Paridis*) have been used as traditional medicine for more than 2,000 years in China [14]. To date, more than 40 commercial drugs and health products have been developed using *Rhizome Paridis* as raw materials [15], with ~1.5 billion USD per year in gross sales [16]. In addition, nearly all species with thick rhizomes are collected for medicinal purposes in Vietnam, Myanmar, Nepal, Bhutan, and India [13-14].

*Paris* is morphologically distinctive in their single whorl of leaves (>3) and solitary apical flower that is 4–15-merous. However, the rhizome, leaf, flower, stamens, ovary, fruit and seeds, which have been widely used to construct classifications, are highly divergent among species [7, 17]. Since the establishment of the genus by Linnaeus [18], it has been subject to numerous critical revisions. Based on rhizome and fruit

morphology, Franchet [19], who established the first infrageneric classification system of *Paris*, placed the species known at that time into two sections: *Euthyra* and *Paris*. Hara [17] described a third section, *Kinugasa*. Instead, Takhtajan [20] recognized these three sections as genera: *Paris s. s.* (= sect. *Paris*), *Daiswa* (= sect. *Euthyra*), and *Kinugasa* (= sect. *Kinugasa*). In the most comprehensive revision, Li [7] divided the genus into two subgenera, *Daiswa* and *Paris*, and eight sections, *Axiparis*, *Dunniana*, *Euthyra*, *Fargesiana*, *Kinugasa*, *Marmorata*, *Paris*, and *Thibetica*. Based on molecular and morphological evidence, Ji et al. [21] suggested an updated classification of Li [7] by combining sections *Dunniana*, *Fargesiana* and *Marmorata* with *Euthyra*.

Several recent studies attempted to reconstruct phylogenetic relationships within *Paris* based on single or multiple DNA loci [21–24]. Due to insufficient sequence variation or limited taxon sampling, these studies did not provide satisfactory resolution or support for infrageneric relationships. As such, the absence of a solid phylogenetic scheme hinders the satisfactory resolution of the long-standing disagreements over classification of *Paris* and limits our understanding of the evolutionary and biogeographic history of this economically important genus.

Plant phylogenetics based on limited sequence regions often suffer from poor resolution and low support, particularly for clades in which rapid diversification or hybridization events have occurred [25–29]. Recently, next-generation sequencing, a technique capable of producing orders of magnitude more data than Sanger sequencing, has been increasingly used for phylogenetic reconstruction [30–35]. This has offered new approaches to resolve recalcitrant relationships in phylogenetically difficult taxa [36–43]. Huang et al. [44] and Yang et al. [45] attempted to apply plastid genomes (plastomes) to resolve phylogenetic relationships within *Paris*. Although the plastome data greatly improve phylogenetic resolution and support, limited taxon sampling prevented them from building a robust overall view of the genus. It is, therefore, necessary to extend sampling size to cover all described sections and even all species, and to use markers with different inheritance patterns to comprehensively understand the evolutionary history of *Paris*.

The biparentally inherited but generally uniparental evolution via gene conversion of nuclear ribosomal DNA sequences (nrDNA) and the non-recombining, mostly maternally inherited plastomes contain a large number of evolutionarily informative variation suitable for phylogenetic analysis [46–50]. Genome skimming via shotgun sequencing of total genomic DNA at relatively low coverage is an efficient approach to recover entire plastomes and nrDNA [47]. Recently, genome skimming has been widely employed to reconstruct the evolutionary relationship at lower taxonomic levels and among closely related species [51–55], as well as to investigate reticulate evolution in diverse plant clades [52, 56–58]. In this study, we generated plastome and nrDNA sequences from all currently recognized *Paris* species using genome skimming method. Based on phylogenomic analyses, we aimed to (1) clarify evolutionary relationships within *Paris*, and (2) explore biogeographic scenarios and historical diversification for the genus.

## Results

## Illumina sequencing and assembly

Low coverage genome sequencing generated per sample 8.57–35.73 million paired-end clean reads (150 bp) (**Additional file 1**: Table S1). Of these,  $7.25 \times 10^4$  to  $2.10 \times 10^6$  and  $4.81 \times 10^3$  to  $4.59 \times 10^4$  were mapped to the reference plastome and nuclear nrDNA, respectively. Based on these data, we assembled complete plastomes and nrDNA for all samples, with the average sequencing depth ranging from 68.72–1998.57 times and 75.15–1136.47 times, respectively.

The *de novo* assembly produced 33 *Paris* plastomes, which exhibited a typical quadripartite structure, with the size varying from 156,139–158,643 bp (**Additional file 2**: Figure S1). *Paris* plastomes are conserved in gene content and arrangement. All plastomes contain 114 genes, including 80 protein-coding genes, 30 tRNA genes, and four plastid rRNA genes (**Additional file 3**: Table S2). Alignment of the plastomes yielded a matrix of 166,726 positions, in which we identified 5,899 variable sites (3.53%) and 3,225 (1.93%) were parsimony informative (Table 1). Also, our *de novo* nrDNA assembly entirely covered 18S, ITS1, 5.8S, ITS2 and 26S regions. The sequence length of *Paris* nrDNA ranged from 5,840–5,859 bp. Alignment of the nrDNA sequences produced 443 variable sites (7.56%), of which 264 (4.50%) were parsimony informative (Table 1).

## Phylogenetic relationships

The standard maximum likelihood (ML) and Bayesian inference (BI) analyses of complete plastomes generated identical tree topologies (Fig. 1). Five highly supported clades (bootstrap percentage, BP=100; posterior probability, PP=1.00) within Melanthiaceae were resolved, which correspond to the five tribes recognized by Zomlefer [3]. Their relationships are congruent with those of previous studies [3, 5, 6, 44, 59]. The monophyly of *Paris* was strongly supported (BP=100, PP=1.00), which was sister to *Trillium* (BP=100, PP=1.00). Within *Paris*, five well-supported clades corresponding to the five sections circumscribed by Ji et al. [21] were recovered. Our results support the successive divergence of the *P.* sections *Paris*, *Kinugasa*, *Thibeticae*, *Axiparis* and *Euthyra*. Most relationships obtained high support except a few terminal species relationships. For instance, the relationship between *P. luquanensis* and *P. marmorata* received weak branch support (BP=57, PP=0.67) in both ML and BI trees.

The incongruence length difference (ILD) test revealed significant discordance ( $p < 0.001$ ) between the nrDNA and plastome datasets. The phylogenetic analysis of nrDNA sequences (Fig. 2) divided tribe Paridae (*Paris* and *Trillium*) species into two clades (BP=100, PP=1.00). Within the first clade (BP=70, PP=0.50), section *Paris* is sister to *Trillium* (BP=59, PP=0.60) and this pair is sister to section *Kinugasa*. Within the second clade (BP=100, PP=1.00), section *Thibeticae* is sister to all species of sections *Axiparis* and *Euthyra*, which are non-monophyletic. In comparing nuclear and plastid topologies, we observed three instances of cytonuclear discordance at different taxonomic levels (Fig. 2). The first is the non-monophyly of *Paris* in the nuclear dataset. The second is monophyly of sections *Axiparis* and *Euthyra*: they are paraphyletic in nrDNA tree but monophyletic in the plastid tree. The third instance of cytonuclear discordance concerns the interspecific relationships within section *Euthyra* (Fig. 2).

## Ancestral area reconstruction, molecular dating, and historical diversification

The Statistical-Dispersal Vicariance Analysis (S-DIVA) (Fig. 1) reconstructed northeastern Asia and northern China (C) as the ancestral area for the most recent common ancestor (MRCA) of *Paris*. It may have undergone a westward or southward dispersal into southwestern China and Himalayas or eastern, central, southern China and northern Indochina [AC (0.50)/BC (0.50)] to evolve the MRCA of thick rhizome clade. Then, a vicariance was inferred to split section *Kinugasa* (Japanese Islands) from remaining taxa (sections *Thibeticae*, *Axiparis*, and *Euthyra*, subtropical East Asia); within the latter, three dispersal and two vicariance events were inferred. Although the S-DIVA analysis failed to reconstruct the ancestral area of the section *Paris*, a dispersal, an extinction, and two vicariance events were inferred in the clade.

The BEAST analyses (Fig. 3) indicated that the divergence between the sister genera, *Paris* and *Trillium*, occurred at 33.94 Mya (95% HPD: 37.84–29.70 Mya). Within *Paris*, the thick and slender rhizome clades diverged from their MRCA at 28.66 Mya (95% HPD: 35.17–20.62 Mya), around the early Oligocene. Diversification within the thick rhizome clade commenced 16.00 Mya (95% HPD: 22.39–7.04 Mya), around the early Miocene, leading to the divergence of the monotypic section *Kinugasa* from the remaining thick-rhizome taxa. Subsequently, the monotypic section *Thibeticae* diverged from the MRCA of sections *Axiparis* and *Euthyra* at 10.08 Mya (95% HPD: 13.51–7.46 Mya), in late Miocene. The split of sections *Axiparis* and *Euthyra* was dated at 7.07 Mya (95% HPD: 9.38–5.12 Mya), around the Miocene/Pliocene boundary. Additionally, the diversification of sections *Paris*, *Axiparis*, and *Euthyra* occurred at 10.93 Mya (95% HPD: 21.14–5.65 Mya), 4.77 Mya (95% HPD: 6.75–2.91 Mya), and 4.59 Mya (95% HPD: 6.27–3.12 Mya), respectively.

The semi-logarithmic lineage through time (LTT) plots analysis (Fig. 4) suggested that the origin of *Paris* was followed by a relatively stable diversification rate, which, however, sharply increased around Miocene/Pliocene boundary. This upward trend was maintained during the Pliocene and the Pleistocene. The Bayesian Analysis of Macroevolutionary Mixtures (BAMM) detected a rate shift in net species diversification in *Paris*, which occurred with the divergence between sect. *Euthyra* and *Axiparis* (Fig. 4).

## Discussion

### Phylogenetic inferences and taxonomic implications

Previous phylogenetic analyses based on a small number of DNA loci or limited taxon sampling failed to robustly reconstruct the backbone of the *Paris* tree [21–24, 44, 45]. Including all currently recognized species, the plastome analysis fully resolved interspecific relationships of *Paris* with strong support at most nodes. Our study further confirms that phylogenetic analysis based on more DNA loci with greatly increased number of phylogenetically informative characters can significantly improve resolution at low taxonomic levels [36, 43, 52, 53, 60].

The plastome-based phylogenies strongly support the monophyly of *Paris* and recovered five strongly supported major clades that correspond to the previously proposed sections by Ji et al. [21]. Within *Paris*,

successive divergence along the spine of the tree of sections *Paris*, *Kinugasa*, *Axiparis*, *Thibeticae*, *Axiparis*, and *Euthyra* was inferred. This divergence pattern can be supported by some morphological characters (Fig. 5). Briefly, the slender rhizome and round more berry-like fruit distinguish section *Paris* from the rest of the sections. Nevertheless, seeds without an enveloping sarcotesta (or aril, presumably a plesiomorphic character) separate sections *Paris* and *Kinugasa* from the rest. Although species of the thick rhizome clade (including sections *Kinugasa*, *Thibeticae*, *Axiparis*, and *Euthyra*) commonly have angular fruits, section *Kinugasa* is distinctive among these in possessing showy white sepals. The third diverging section *Thibeticae* is similar to *Euthyra* in having dehiscent capsules, but its seed morphology (with an incomplete aril) is analogous to that of *Axiparis*. Therefore, the plastid tree is highly congruent with the morphologically based classification of Ji et al. [21].

The plastid tree provides valuable insights for resolving the long-standing disagreements in classification of *Paris*. All *Paris* species share the morphological synapomorphies of flowers and leaves being 4–15-merous compared with the trimerous condition of *Trillium* (Fig. 5), and monophyly of both genera was fully supported, making it reasonable to recognize *Paris* as a single genus [7, 17, 21] rather than dividing it into three genera (*Daiswa*, *Kinugasa*, and *Paris s. s.*) [20]. Also, the new tree further supports the taxonomic treatment of Ji et al. [21] by combining sections *Dunniana*, *Fargesiana* and *Marmorata* with *Euthyra*. Given its economic importance, resolution of the long-standing taxonomic disputes will be conducive to exploration and protection of *Paris* species.

The current taxonomy of *P. polyphylla* with five varieties is not supported by either the plastid or nrDNA results. The varieties should be probably recognized as distinct species, but it is also likely, given the cytonuclear discordance observed for these accessions (Fig. 2) that hybridization may be involved in their origins. Further study of this group with more appropriate population genetic and cytological techniques is warranted.

### **Cytonuclear discordance**

Similar to the previous study of Ji et al. [21], we detected discordance between the nrDNA and plastid trees for *Paris* at both deep and shallow nodes. Cytonuclear incongruence is a fairly common phenomenon in plant phylogenetics [25, 56, 61–68]. In most cases, the nuclear tree is more congruent with morphological taxonomy [43, 56, 61, 62, 64, 67–70], and such incongruence can be mainly attributed to incomplete sorting of cytoplasmic polymorphisms or introgression of the cytoplasmic genome from one species into the nuclear background of another by hybridization [25, 63, 68, 71, 72]. However, in our study the plastid tree of *Paris* is largely consistent with morphological evidence, suggesting nrDNA introgression without cytoplasmic gene flow could be responsible for the discordance detected [68, 71–74].

Given that the discordance observed in *Paris* was likely due to phenomena affecting the nrDNA tree, which involved rapid gene conversion of one parental copy [75–80], the phylogenetic relationships recovered by this dataset may not be representative of that obtained with other parts of the nuclear genome not subject to gene conversion. Nonetheless, these results may provide useful information about

past hybridization events that otherwise may not be the majority pattern in the nuclear genome. Failure to recover *Paris* as a monophyletic in the nrDNA tree (Fig. 2) suggests there could have been hybridization between section *Paris* and *Trillium* after section *Paris* split from the rest of the genus. It is noteworthy that the largest eukaryotic genome, that *P. japonica* [81, 82], was supposed to be an allo-octaploid between *Paris* and *Trillium* according to previous cytological investigations [83]. This hybrid hypothesis is supported by its position as sister to section *Paris* and *Trillium* in the nrDNA tree. Likewise, the non-monophyly of sections *Axiparis* and *Euthyra* observed in the nrDNA tree (Fig. 2) can also be attributed to ancient hybridization. The sister relationship of *P. luquanensis*/*P. mairei* (section *Euthyra*) and the clade of four species of section *Axiparis* (*P. dulongensis*, *P. forrestii*, *P. rugosa* and *P. tengchongensis*) suggest that hybridization could have occurred between the ancestors of these taxa. Additionally, extensive discordance among species of section *Euthyra* (Fig. 2) supports the conclusions of the previous study that natural hybridization between species of section *Euthyra* is likely if the pollinators are the same, but little is known about this aspect of the biology of *Paris*. Experimental manual outcrossing has been effective between most of these species [84]. Interspecific hybridization is the likely cause of the cytonuclear discordance observed between species in the section. Additionally, as mentioned above, there is discordance for the positions of the varieties of *P. polyphylla*, suggesting that hybridization may also have played a role in their origins.

It is notable that the cytonuclear discordance detected in this study merely reflects conflict between plastomes and nrDNA datasets, which are substantially two linkage groups of plastid and ribosomal genes and only represent at best two large single-locus DNA regions. Whether there is nuclear genome-wide and plastome discordance is a gigantic leap with such a limited dataset. To further verify this, large numbers of unlinked nuclear loci generated by restricted site-associated DNA sequencing (RAD-seq) or even whole nuclear genome sequencing are likely to be required.

### **Biogeography and lineage diversification**

Because strong cytonuclear discordance was detected in *Paris* and the plastid tree agrees well with morphologically based classification, we address the biogeography and historical diversification of *Paris* based on the plastid dataset. The S-DIVA analysis recovered northeastern Asia and northern China as the ancestral area of *Paris*. Associated with a dispersal and a vicariance event (Fig. 1), the crown node of *Paris* was dated at 28.66 Mya (Fig. 3), in the early Oligocene, when the global climatic deterioration [85] led to the expansion of vegetation adapted to drier and colder climates in large parts of Eurasia [86]. Therefore, early divergence of *Paris* may have been driven by these events. Also, S-DIVA analysis revealed that the divergence of the Japanese endemic species, *P. japonica* and *P. tetraphylla* was triggered by two independent vicariance events (Fig. 1). Their divergence times of 16.00 Mya (*P. japonica*) and 10.93 Mya (*P. tetraphylla*), in the Miocene corresponds to the opening of the Japan Sea, which separated the Japan islands from the continental East Asia [87].

In the thick rhizome clade, the S-DIVA analysis inferred three dispersal events (Fig. 1), which were dated at 10.08, 7.07 and 4.59 Mya (Fig. 3), respectively. Neocene climatic change might play an essential role in

triggering these events. The onset of the Asian monsoon around the Oligocene/Miocene transition created a connection between forests from low to high latitudes of East Asia [88]. The enhancement of Asian summer monsoon since the late Miocene established a humid climate in subtropical East Asia [89–91], and caused a significant expansion of forests in East Asia [88, 92]. These climatic and environmental shifts would create favorable habitats that facilitated the dispersal and divergence of sections *Thibeticae*, *Axiparis* and *Euthyra* in subtropical East Asia. Also, along with the expansion of forests in high latitudes of East Asia, the MRCA of *P. quadrifolia* and *P. incompleta* may have migrated into Europe.

Both LTT and BAMM analyses revealed that clade diversification within *Paris* abruptly accelerated around the Miocene/Pliocene boundary, which could be driven by the further strengthening of monsoonal climate in the summer and the initiation of the two distinct monsoon regimes that have gradually become established in subtropical East Asia since the late Miocene [88, 92, 93]. From then on, eastern, central and southern China and northern Indochina have been primarily governed by Pacific monsoon, whereas southwestern China and the Himalayas have been mainly affected by Indian monsoon [92–95]. The S-DIVA analysis (Fig. 1) and molecular dating (Fig. 3) showed that vicariance events occurred independently in sections *Axiparis* (4.77 Mya) and *Euthyra* (4.28 Mya) in the two regions mentioned above. This implies profound ecological heterogeneity resulting from climate differentiation may have driven significant allopatric speciation in the two regions [96–98]. In addition, the LTT and BAMM analyses (Fig. 4) revealed that most extensive divergence in *Paris*, which was responsible for appearance of more than half of the extant taxa, took place in the Pliocene and the Pleistocene. It is believed that the Qinghai-Tibet Plateau (QTP) rose dramatically from the Late Miocene (*ca.* 10~8 Ma) to the early Pliocene (*ca.* 3.6 Ma) [99, 100], which dramatically modified global climate [94, 101] and thereby profoundly influenced biological processes, such as species range expansion/contraction and vicariance, in East Asia [102]. During the Pleistocene, there were at least four major glaciations in East Asia [103], and these probably created significant isolation and diverse habitats in East Asia [104, 105]. Such complex geological, ecological, and environmental heterogeneity is expected to have driven diversification of a wide spectrum of plant clade [104, 106–109] and would also have triggered vicariance and facilitated a species radiative in *Paris*.

A negative correlation between genus-level diversity and the genus-average genome size was observed in plants [110, 111]. Knight *et al.* [111] proposed the large genome constraint hypothesis, which states that plant taxa with large size genomes diversify more slowly. Subsequently, Suda *et al.* [112] found that many island clades of Macaronesian angiosperms that underwent adaptive radiations have small genome sizes, and assumed that rapid diversification is more likely to happen in angiosperms with small genomes size. It is noteworthy that *Paris* is fairly distinctive in angiosperms in possessing large genomes. The minimum documented genome size in the genus (*P. verticillata*, 1C = 30.52 Gb) is much larger than the mean genome size (1C = 5.7 Gb) of angiosperms [113, 114]. Moreover, the known largest eukaryotic genome, that of *P. japonica*, 1C = 148.88 Gb, belongs to *Paris* [81, 82]. In this study, we found that *Paris* may have undergone a species radiation since the Miocene/Pliocene boundary (Fig. 4), which is not consistent with prediction that large genome size could limit speciation [111, 112]. It also has to be admitted that although there is sharp rise in the lineage diversification, the total number of species

involved is not large in comparison to other radiations, for instance, *Dianthus* in the Mediterranean [115], and *Aizoaceae* in South Africa [116]. The generality of the large genome constraint hypothesis needs to be further evaluated, although the increased lineage diversification detected here in *Paris* does not pose a major contradiction to it.

## Conclusions

This study represents a comprehensive phylogenetic investigation, an economically important but taxonomically difficult genus, by sampling all currently recognized species in the genus. The analyses of complete plastome and nrDNA sequences reconstructed a robust phylogeny, and provided implications for clarifying some of the long-standing taxonomic debates in *Paris*. We also identified significant conflicts between plastid and nuclear datasets. This cytonuclear discordance observed in *Paris* may have been caused by ancient and recent hybridizations. Ancestral area reconstruction indicated that *Paris* may have originated in northeastern Asia and northern China, and has experienced multiple dispersal and vicariance events during its diversification. Based on phylogenetic framework and molecular dating, we propose that the climatic and geological changes since the Miocene played essential roles in triggering range formation and clade diversification in *Paris*. Our findings provide important insights for elucidating the evolutionary history of *Paris*, and will be conducive to exploration and protection of *Paris* species.

## Methods

### Plant sampling, DNA extraction and Illumina sequencing

We sampled 33 accessions to represent all 29 species and five varieties recognized by Li [7] and those described since then [9, 84, 117–119]. The original sources of the plant materials used in this study and voucher information are presented in **Additional file 4** (Table S3). The voucher specimens were identified by Dr. Yunheng Ji. Genomic DNA was extracted from *ca.* 20 mg silica gel dried leaves using the CTAB (cetyltrimethylammonium bromide) method [120]. Approximately 5 µg of purified genomic DNA was sheared by sonication. Paired-end libraries with an average insert size 350 bp were prepared using a TruSeq DNA Sample Prep Kit (Illumina, Inc., USA) according to the manufacturer's protocol. The libraries were paired-end sequenced on the Illumina HiSeq 2000 platform. Raw reads were filtered to remove adaptors and low quality reads using NGS QC Toolkit [121], by setting the cut-off value for percentage of read length to 80 and PHRED quality scores at 30.

### Assembly and gene annotation

The complete plastome sequence of *Paris quadrifolia* (GenBank Accession: KM067394) was used as the reference for assembling the newly sequenced *Paris* plastomes. The plastid contigs were organized according to the references and connected with overlapping terminal sequences to yield the complete plastomes in Bowtie v2.2.6 [122] using the default parameters. Plastomes were annotated with the Dual Organellar Genome Annotator database [123]. Start and stop codons and intron/exon boundaries for

protein-coding genes were checked manually. Annotated tRNA genes were further verified by tRNAscan-SE 1.21 [124] with the default parameters. The boundary of the large-single copy (LSC), small-single copy (SSC), and inverted-repeat (IR) regions for each plastome were visually examined and manually adjusted according to those of the reference plastome in Geneious V10.2 [125].

For nrDNA sequence assembly, we first excluded all plastid-like reads. Based on remaining reads, *de novo* assemblies were performed using the complete nrDNA sequence (including 26S, 18S and 5.8S ribosomal RNA genes and the internal transcribed spacers) of *Lillium tsingtauense* (GenBank Accession: KM117263) as reference. The external transcribed spacers in *Paris* species possesses too many repeat sequences and inversions that may make the assembly inaccurate, we therefore did not assemble the region. Contigs mapping to reference nrDNA were assembled using the processes described above. The nuclear ribosomal RNA genes and their boundaries with ITS regions were annotated and defined by comparison with the reference in Geneious V10.2 [125].

### Phylogenetic analysis

To investigate phylogenetic placement of *Paris* within Melanthiaceae, 15 other plastomes representing the five tribes recognized in Melanthiaceae (**Additional file 5**: Table S4) were integrated with the 33 newly sequenced *Paris* plastomes in the final analysis. Furthermore, 33 *Paris* nrDNA and five rDNA sequences from *Veratrum* (1 accession), *Ypsilandra* (1 accession) and *Trillium* (3 accessions) of Melanthiaceae were incorporated into a nuclear dataset (**Additional file 4**: Table S3). *Campynema lineare* and *Veratrum taliense* were used to root the plastid and nuclear trees, respectively, according to previous studies [3, 5, 6, 59]. Alignment of plastid and nrDNA sequences were performed using MAFFT [126] integrated in Geneious v.10.2 [125], and manually edited if necessary. The most appropriated model of sequence substitution for plastomes (GTR+G) and nrDNA sequences (GTR+I+G) was selected using Modeltest v3.7 [127] with the Akaike information criterion [128]. We considered the whole plastome as a single inherited unit. Next, we confirmed the same model for both the small and large single copy regions and the inverted repeats using PartitionFinder v. 2.1.1 [129]. Conflict between plastid and nuclear datasets was examined using the incongruence length difference (ILD) test [130] implemented in PAUP\* 4.0b10 [131] for 1,000 replicates.

Phylogenetic analyses were carried out using both ML and BI methods. ML analyses were conducted using RAxML-HPC BlackBox v8.1.24 [132] with 1,000 replicates of rapid bootstrapping. The BI analyses were performed using MRBAYES v.3.1.2 [133]. Runs for each dataset began with a random starting tree for one million generations with sampling at every 100 generations. An initial 25% of the sampled trees were discarded. The posterior probability values were determined from the remaining trees. Stationarity was considered to be reached when the average standard deviation of the split frequencies was < 0.01.

### Molecular dating and diversification rate estimate

It is notable that no fossils have been identified for Melanthiaceae and its close relatives. A previous study that used 17 fossils across the monocots and major clades of angiosperms suggested that the

crown age of Melanthiaceae was approximately 84.8 Mya, while the clades Parideae-Xerophyllideae and Chionographideae-Heloniadeae diverged approximately 74 Mya, and the tribes Parideae and Xerophyllideae split approximately 52.3 Mya [134]. We used these times to calibrate the phylogenetic tree (Fig. 3). Molecular dating was performed in BEAST v.2.4.7 [135]. The BEAST analyses were run under the uncorrelated lognormal relaxed clock approach with a Yule tree prior. Markov Chain Monte Carlo chains were run for 10,000,000 were run with sampling every 1000 generations. The stationarity of the chains and convergence of BEAST analyses was monitored by Tracer v. 1.5.

The diversification rate change over time was inferred using the semi-logarithmic lineage through time (LTT) plot approach. The consensus chronogram inferred from the results of molecular dating was computed by APE v.5.3 [136] within an R environment [137]. We further examined potential shifts in net diversification rate in *Paris* based on the time calibrated maximum clade credibility tree (with the highest posterior probability) from BEAST using the Bayesian Analysis of Macroevolutionary Mixtures (BAMM) [138].

### **Ancestral area reconstruction**

For biogeographic reconstructions, *Paris* species were assigned to four areas based on their current distributions: A) southwestern China and Himalayas, B) eastern, central, southern China and northern Indochina, C) northeastern Asia and northern China, and D) Europe and the Caucasus. Ancestral distributions of *Paris* were reconstructed by statistical dispersal-vicariance analysis (S-DIVA) [139] implemented in RASP 4.0 [140]. The condensed tree and 4,000 post burn-in Bayesian trees from BI analysis were used as input trees. The random tree was defined as 1,000 and other parameters were set to their defaults.

## **Abbreviations**

USD: United States dollar; DNA: Deoxyribonucleic acid; nrDNA: nuclear ribosomal DNA sequences; bp: Base pair; tRNA: Transfer RNA; rRNA: Ribosomal RNA; ITS: internal transcribed spacer of nuclear ribosomal DNA; ML: Maximum Likelihood; BI: Bayesian Inference; BP: Bootstrap percentage; PP: Posterior probability; ILD: incongruence length difference; S-DIVA: the Statistical-Dispersal Vicariance Analysis; MRCA: most recent common ancestor; HPD: Highest posterior density; Mya: Million years ago; LTT: semi-logarithmic lineage through time; QTP: Qinghai-Tibet Plateau; KUN: herbarium of Kunming Institute of Botany, Chinese Academy of Sciences; CTAB: Cetyl trimethylammonium bromide; LSC: Large single-copy; SSC: Small single copy; IR: Inverted repeat; BS: Bootstrap; MCMC: Markov Chain Monte Carlo.

## **Declarations**

### **Ethics approval and consent to participate**

Collection of all samples completely complies with national and local legislation permission. Plant samples used in the study were not collected from national park or natural reserve. According to national

and local legislation, no specific permission was required for collecting these plants.

### **Consent for publication**

Not applicable.

### **Availability of data and materials**

The sequences generated in this study are available at GenBank (accession numbers are presented in Table S3). The manually adjusted alignment of sequences are deposited in Treebase (<http://purl.org/phylo/treebase/phylows/study/TB2:S25166>).

### **Competing interests**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### **Funding**

The authors would like to thank the financial support from the Major Program of National Natural Science Foundation of China (31590823), the National Natural Science Foundation of China (31872673), the NSFC-Joint Foundation of Yunnan Province (U1802287), a grant of the Large-scale Scientific Facilities of the Chinese Academy of Sciences (No. 2017-LSF-GBOWS-02), and, and the Major Science and Technology Projects of Yunnan Science and Technology Plan (2019ZF011-2) in design of the study, and collection, analysis and interpretation of data, as well as in writing the manuscript.

### **Authors' contributions**

YJ, JBY, and TSY conceived the study; YJ, LY, and JBY collected the data; LY, CL, ZY, and JY analyzed the data; YJ wrote the manuscript; MWC and TSY discussed the results and revised the manuscript. All authors have read and approved the manuscript

### **Acknowledgements**

We thank Royal Botanic Garden (Edinburgh), Heng Li, Bing Liu, and Pascal Bruggeman for providing some materials in this study; Sirong Yi, Haicheng An, Qiliang Gan, Ren Zhao, Tingzhou Zhao, Chengjin Yang, Guohua Zhou, and Zhangming Wang for their help with field work.

## **References**

1. Lindley J. The vegetable kingdom, 3<sup>rd</sup> edition. London: Bradely and Evans; 1846.
2. Angiosperm Phylogeny Group. An ordinal classification for the families of flowering plants. *Ann Missouri Bot Gard.* 1998;85:531–553. doi: 10.2307/2992015.

3. Zomlefer WB, Williams NH, Whitten WM, Judd WS. Generic circumscription and relationships in the tribe Melanthieae (Liliales, Melanthiaceae), with emphasis on *Zigadenus*: evidence from ITS and *trnL-F* sequence data. *Am J Bot.* 2001;88:1657–1669. doi: 10.2307/3558411.
4. Zomlefer WB, Judd WS, Whitten WM, Williams NH. A synopsis of Melanthiaceae (Liliales), with focus on character evolution in tribe Melanthieae. In: Columbus JT, Friar EA, Porter JM, Prince LM, Simpson MG, editors. *Monocots: Comparative biology and evolution (excluding Poales)*. California: Rancho Santa Ana Botanic Garden; 2006. p. 564–576.
5. Kim JS, Hong JK, Chase MW, Fay MF, Kim JH. Familial relationships of the monocot order Liliales based on a molecular phylogenetic analysis using four plastid loci: *matK*, *rbcl*, *atpB* and *atpF-H*. *Bot J Linn Soc.* 2013;172:5–21. doi: 10.1111/boj.12039.
6. Kim S, Kim JS, Chase MW, Chase MW, Fay MF, Kim J. Molecular phylogenetic relationships of Melanthiaceae (Liliales) based on plastid DNA sequences. *Bot J Linn Soc.* 2016;181:567-584. doi:10.1111/boj.12405.
7. Li, H. The genus *Paris* (Trilliaceae). Beijing: Science Press; 1998.
8. Ji Y, Zhou Z, Li H. Four new synonyms in the genus *Paris* (Trilliaceae). *Acta Phytotaxon Sin.* 2007;45:388–390. doi: 10.1360/aps06163.
9. Yang J, Wang YH, Li H. *Paris qiliangiana* (Melanthiaceae), a new species from Hubei, China. *Phytotaxa.* 2017;329:193–196.
10. Liang SY, Soukup VG. *Paris* L. In: Wu ZY, Raven PH, editors. *Flora of China*. Beijing: Science Press and Missouri Botanical Garden Press; 2000. p. 88–95.
11. Long CL, Li H, Ouyang Z, Yang X, Li Q, Trangmar B. Strategies for agrobiodiversity conservation and promotion: a case from Yunnan, China. *Biodiversity Conserv.* 2003;12:1145–1156. doi:10.1023/a:1023085922265.
12. He J, Zhang S, Wang H, Chen CX, Chen SF. Advances in studies on and uses of *Paris polyphylla* var. *yunnanensis* (Trilliaceae). *Acta Bot Yunnan.* 2006;28:271–276.
13. Cunningham AB, Brinckmann JA, Bi YF, Pei SJ, Schippmann U, Luo P. *Paris* in the spring: a review of the trade, conservation and opportunities in the shift from wild harvest to cultivation of *Paris polyphylla*(Trilliaceae). *J Ethnopharmacol.* 2018;222:208–216. doi: 10.1016/j.jep.2018.04.048.
14. Li H, Su B, Yang Y, Zhang Z. An assessment on the rarely medical *Paris* plants in China with exploring the future development of its plantation. *J West China For Sci.* 2015;44:1–6.
15. Wang YH, Niu HM, Zhang ZY, Hu XY, Li H. Medicinal values and their chemical bases of *Paris*. *China J Chin Mater Med.* 2015;40:833–839.
16. Huang LQ, Xiao PG, Wang YY. Investigation on resources of rare and endangered medicinal plants in China, first edn. Shanghai: Science & Technology Press; 2012.
17. Hara, H. Variation in *Paris polyphylla* Smith, with reference to other Asiatic species. *J Fac Sci Univ Tokyo.* 1969;10:141–180.
18. Linnaeus C. *Species plantarum*. Salvius: Stockholm; 1753.

19. Franchet A. Monographie du genre *Paris*. Mem Soc Philom Centen. 1888;24:267–291.
20. Takhtajan A. A revision of *Daiswa* (Trilliaceae). Brittonia. 1983;35:255–270.
21. Ji Y, Fritsch PW, Li H, Xiao TJ, Zhou ZK. Phylogeny and classification of *Paris* (Melanthiaceae) inferred from DNA sequence data. Ann Bot. 2006;98:245–256. doi: 10.1046/j.1365-2885.2002.00426.x .
22. Kato H, Terauchi R, Utech FH, Kawano S. Molecular systematics of the Trilliaceae *sensu lato* as inferred from *rbcL* sequence data. Mol Phylogenet Evol. 1995;4:184–193. doi: 10.1006/mpev.1995.1018.
23. Osaloo SK, Kawano S. Molecular systematics of Trilliaceae II. Phylogenetic analyses of *Trillium* and its allies using sequences of *rbcL* and *matK* genes of cpDNA and internal transcribed spacers of 18S–26S nrDNA. Plant Species Bio. 1999;14:75–94. doi:10.1046 /j.1442-1984.1999.00009.x.
24. Farmer SB, Schilling EE. Phylogenetic analyses of Trilliaceae based on morphological and molecular data. Syst Bot. 2002;27:674–692. doi: 10.2307/3093915.
25. Rieseberg LH, Soltis DE. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol Trends Plant. 1991;5:64–84. doi: 10.1007/BF00021248.
26. Rokas A, Carroll SB. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol Biol Evol. 2005;22:1337–1344. doi: 10.1093/molbev/msi121.
27. Whitfield JB, Lockhart PJ. Deciphering ancient rapid radiations. Trends Ecol Evol. 2007;22:258–265. doi: 10.1016/j.tree.2007.01.012.
28. Philippe H, Brinkmann H, Lavrov DV, Littlewood DTJ, Manuel M, Wörheide G, et al. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 2011; 9: e1000602. doi:10.1371/journal.pbio.1000602.
29. Sun M, Soltis DE, Soltis PS, Zhu X, Burleigh JG, Chen Z. Deep phylogenetic incongruence in the angiosperm Rosidae clade. Mol Phylogenet Evol. 2015;83:156–166. doi: 10.1016/j.ympev.2014.11.003.
30. Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, et al. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. PNAS. 2007;104:19369–19374. doi: 10.1073/pnas.0709121104.
31. Moore MJ, Bell CD, Soltis PS, Soltis DE. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc Natl Acad Sci. 2007;104:19363–19368. doi:10.1073/pnas.0708072104.
32. Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc Natl Acad Sci. 2010;107:4623–4628. doi: 10.1073/pnas.0907801107.
33. Philippe H, Derelle R, Lopez P, Pick K, Borchiellini C, Boury-Esnault N, et al. Phylogenomics revives traditional views on deep animal relationships. Curr Biol. 2009;19:706–712. doi: 10.1016/j.cub.2009.02.052.

34. Zhou X, Xu S, Xu J, Chen B, Zhou K, Yang G. Phylogenomic analysis resolves the interordinal relationships and rapid diversification of the Laurasiatherian mammals. *Syst Biol.* 2012;61:150–164. doi:10.2307/41515183.
35. Mckain MR, Johnson MG, Uribe-Convers S, Eaton D, Yang Y. Practical considerations for plant phylogenomics. *Appl Plant Sci.* 2018;6:e1028. doi: 10.1002/aps3.1038.
36. Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology.* 2009;7:84. doi: 10.1186/1741-7007-7-84.
37. Barrett CF, Specht CD, Leebens-Mack J, Stevenson DW, Zomlefer WB, Davis JI. Resolving ancient radiations: can complete plastid gene sets elucidate deep relationships among the tropical gingers (Zingiberales)? *Ann Bot.* 2014;113:119–133. doi: 10.1093/aob/mct264.
38. Ma PF, Zhang YX, Zeng CX, Guo ZH, Li DZ. Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe *Arundinarieae* (Poaceae). *Syst Biol.* 2014;63:933–950. doi: 10.1093/sysbio/syu054.
39. Stull GW, Dunod SR, Soltis DE, Soltis PS. Resolving basal lamiid phylogeny and the circumscription of Icacinaceae with a plastome-scale data set. *Am J Bot.* 2015;102:1794–1813. doi: 10.3732/ajb.1500298.
40. Attigala L, Wysocki WP, Duvall MR, Clark LG. Phylogenetic estimation and morphological evolution of *Arundinarieae* (Bambusoideae: Poaceae) based on plastome phylogenomic analysis. *Mol Phylogenet Evol.* 2016;101:111–121. doi: 10.1016/j.ympev.2016.05.008.
41. Zhang SD, Jin JJ, Chen SY, Chase MW, Soltis DE, Li HT, et al. Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytol.* 2017;214:1355–1367. doi: 10.1111/nph.14461.
42. Carlsen MM, Fér T, Schmickl R, Leong-Škorničková J, Newman M, Kress WJ. Resolving the rapid plant radiation of early diverging lineages in the tropical Zingiberales: Pushing the limits of genomic data. *Mol Phylogenet Evol.* 2018;128:55–68. doi: 10.1016/j.ympev.2018.07.020.
43. Lin HY, Hao YJ, Li JH, Fu CX, Soltis PS, Soltis DE, et al. Phylogenomic conflict resulting from ancient introgression following species diversification in *Stewartia s.l.* (Theaceae). *Mol Phylogenet Evol.* 2019;135:1–11. doi: org/10.1016/j.ympev.2019.02.018.
44. Huang Y, Li X, Yang Z, Yang C, Yang J, Ji Y. Analysis of complete chloroplast genome sequences improves phylogenetic resolution of *Paris* (Melanthiaceae). *Front Plant Sci.* 2016;7:1797. doi: 10.3389/fpls.2016.01797.
45. Yang L, Yang Z, Liu C, He Z, Zhang Z, Yang J, et al. Chloroplast phylogenomic analysis provides insights into the evolution of the largest eukaryotic genome holder, *Paris japonica* (Melanthiaceae). *BMC Plant Biol.* 2019;19:293. doi: 10.1186/s12870-019-1879-7.
46. Linder CR, Goertzen LR, Heuvel BV, Francisco-Ortega J, Jansen RK. The complete external transcribed spacer of 18S-26S rDNA: amplification and phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. *Mol Phylogenet Evol.* 2000;14:285–303. doi: 10.1006/mpev.1999.0706.

47. Straub SC, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *Am J Bot.* 2012;99:349–364. doi: 10.3732/ajb.1100335.
48. Bock DG, Kane NC, Ebert DP, Rieseberg LH. Genome skimming reveals the origin of the Jerusalem artichoke tuber crop species: neither from Jerusalem nor an artichoke. *New Phytol.* 2014;201:1021–1030. doi: 10.1111/nph.12560.
49. Hollingsworth PM, Li DZ, Michelle VDB, Twyford AD. Telling plant species apart with DNA: from barcodes to genomes. *Philos Trans R Soc B.* 2016;371:20150338. doi: 10.1098/rstb.2015.0338.
50. Zeng CX, Hollingsworth PM, Yang J, He ZS, Zhang ZR, Li DZ, et al. Genome skimming herbarium specimens for DNA barcoding and phylogenomics. *Plant Methods.* 2018;14:43. doi: org/10.1186/s13007-018-0300-0.
51. Simpson MG, Guilliams CM, Hasenstab-Lehman KE, Mabry ME, Ripma L. Phylogeny of the popcorn flowers: use of genome skimming to evaluate monophyly and interrelationships in subtribe *Amsinckiinae* (Boraginaceae). *Taxon.* 2017;66:1406-1420. doi: 10.12705/666.8.
52. Vargas OM, Ortiz EM, Simpson BB. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytol.* 2017;214:1736–1750. doi: 10.1111/nph.14530.
53. Uribe-Convers S, Carlsen MM, Lagomarsino LP, Muchhala N. Phylogenetic relationships of *Burmeistera* (Campanulaceae: Lobelioideae): combining whole plastome with targeted loci data in a recent radiation. *Mol Phylogenet Evol.* 2017;107:551-563. doi: org/10.1016/j.ympbev.2016.12.011.
54. Fonseca LHM, Lúcia GL. Combining high-throughput sequencing and targeted loci data to infer the phylogeny of the "*Adenocalymma-Neojobertia*" clade (Bignoniaceae, Bignoniaceae). *Mol Phylogenet Evol.* 2018;123:1–15. doi: 10.1016/j.ympbev.2018.01.023.
55. Heckenhauer J, Paun O, Chase MW, Ashton PS, Kamariah AS, Samuel R. Molecular phylogenomics of the tribe *Shoreeae* (Dipterocarpaceae) using whole plastid genomes. *Ann Bot.* 2019;20:1–9. doi: 10.1093/aob/mcy220.
56. Bruun-Lund S, Clement WL, Kjellberg F, Rønsted N. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). *Mol Phylogenet Evol.* 2017;109:93–104. doi: 10.1016/j.ympbev.2016.12.031.
57. Gitzendanner MA, Yang Y, Wickett NJ, Mckain M, Beaulieu JM. Methods for exploring the plant tree of life. *Appl Plant Sci.* 2018;6:e1309. doi: 10.1002/aps3.1039.
58. Liu SH, Edwards CE, Hoch PC, Raven PH, Barber JC. Genome skimming provides new insight into the relationships in *Ludwigia* section *Macrocarpon*, a polyploid complex. *Am J Bot.* 2018;105:875-887. doi: 10.1002/ajb2.1086.
59. Kim C, Kim SC, Kim JH. Historical biogeography of Melanthiaceae: a case of out-of-North America through the Bering land bridge. *Front Plant Sci.* 2019;10:396. doi: org/10.3389/fpls.2019.00396.
60. Ma ZY, Wen J, Ickert-Bond SM, Nie ZL, Chen LQ, Liu XQ. Phylogenomics, biogeography and adaptive radiation of grapes. *Mol Phylogenet Evol.* 2018;129:258-267. doi: org/10.1016/j.ympbev.2018.08.021.

61. Soltis DE, Kuzoff RK. Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae): evidence of chloroplast capture and paraphyly. *Evolution*. 1995;49:727–742. doi: 10.2307/2410326.
62. Soltis DE, Johnson LA, Looney C. Discordance between ITS and chloroplast topologies in the *Boykinia* group (Saxifragaceae). *Syst Bot*. 1996;21:169–176. Doi: 10.2307/2419746.
63. Philippe H, Delsuc F, Brinkmann H, Lartillot N. Phylogenomics. *Annu Rev Ecol Evol Syst*. 2005;36:541–562. doi: 10.1146/annurev.ecolsys.35.112202.130205.
64. Wendel J, Doyle J. Phylogenetic incongruence, window into genome history and evolution. In: Soltis D, Soltis PJD, editors. *Molecular Systematics of Plants II DNA Sequencing*. New York: Kluwer Academic Press; 1998. p. 265–296.
65. Wiens JJ. Combining data sets with different phylogenetic histories. *Syst Biol*. 1998;47:568-581. doi: 10.1080/106351598260581.
66. Quicke DL, Jones OR, Epstein DR. Correcting the problem of false incongruence due to noise imbalance in the incongruence length difference test. *Syst Biol*. 2007;56:496–503. doi: 10.1080/10635150701429974.
67. Xiang XG, Schuiteman A, Li DZ, Huang WC, Chung SW, Li JW, et al. Molecular systematics of *Dendrobium* (Orchidaceae, Dendrobieae) from mainland Asia based on plastid and nuclear sequences. *Mol Phylogenet Evol*. 2013;69:950-960. doi: 10.1016/j.ympev.2013.06.009.
68. Folk RA, Mandel JR, Freudenstein JV. Ancestral gene flow and parallel organellar genome capture result in extreme phylogenomic discord in a lineage of angiosperms. *Syst Biol*. 2017;66:320-337. doi: 10.1093/sysbio/syw083.
69. Wendel JF, Doyle JJ. Phylogenetic incongruence: window into genome history and speciation. In: Soltis PS, Soltis DE, Doyle JJ, editors. *Molecular systematics of plants*. New York: Chapman and Hall; 1998. p.265-296.
70. Renoult JP, Kjellberg F, Grout C, Santoni S, Khadari B. Cyto-nuclear discordance in the phylogeny of *Ficus* section *Galoglychia* and host shifts in plant-pollinator associations. *BMC Evol Biol*. 2009;9:248. doi:10.1186/1471-2148-9-248.
71. Rieseberg LH, Wendel JF. Introgression and its consequences in plants. In: Harrison RG, editor. *Hybrid zones and the evolutionary process*. New York: Oxford University Press; 1993. p. 70–
72. Seehausen O. Hybridization and adaptive radiation. *Trends Ecol Evol*. 2004;19:198–207. doi: 10.1016/j.tree.2004.01.003.
73. Arnold ML, Buckner CM, Robinson JJ. Pollen mediated introgression and hybrid speciation in Louisiana irises. *Proc Natl Acad Sci USA*. 1991;88:1398–1402. doi: 10.2307/2356002.
74. Setoguchi H, Watanabe I. Intersectional gene flow between insular endemics of *Ilex* (Aquifoliaceae) on the Bonin Islands and the Ryukyu Islands. *Am J Bot*. 2000;87:793–810. doi: 10.2307/2656887.
75. Hillis DM, Moritz C, Porter CA, Baker RJ. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. *Science*. 1991;251:208–310. doi: 10.1126/science.1987647.

76. Wendel JF, Schnabel A, Seelanan T. An unusual ribosomal DNA-sequence from *Gossypium gossypoides* reveals ancient, cryptic intergenomic introgression. *Mol Phylogenet Evol.* 1995a;4:298–313.
77. Wendel JF, Schnabel A, Seelanan T. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc Natl Acad Sci USA.* 1995b;92:280–284.
78. Lim KY, Kovarik A, Matyasek R, Bezdek M, Lichtenstein CP, Leitch AR, et al. Gene conversion of ribosomal DNA in *Nicotiana tabacum* is associated with undermethylated, decondensed and probably active gene units. *Chromosome.* 2000;109:161–172 doi: 10.1007/s004120050424.
79. Kovarik A, Pires JC, Leitch AR, Lim KY, Sherwood AM, Matyasek R, et al. Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics.* 2005;169:931–944. doi: 10.1534/genetics.104.032839.
80. Lunerova J, Renny-Byfield S, Matyasek R, Leitch A, Kovarik A. Concerted evolution rapidly eliminates sequence variation in rDNA coding regions but not in intergenic spacers in *Nicotiana tabacum* allotetraploid. *Plant Syst Evol.* 2017;303:1043–1060. doi: 10.1007/s00606-017-1442-7.
81. Pellicer J, Fay MF, Leitch IJ. The largest eukaryotic genome of them all? *Bot J Linn Soc.* 2010;164:10–15. doi: 10.1111/j.1095-8339.2010.01072.x.
82. Dodsworth S, Leitch AR, Leitch IJ. Genome size diversity in angiosperms and its influence on gene space. *Curr Opin Genet Dev.* 2015;35:73–78. doi: 10.1016/j.gde.2015.10.006.
83. Haga T. Chromosome complement of *Kinugasa japonica* with special reference to its origin and behavior. *Cytologia.* 1938;8:137–141. doi: 10.1508/cytologia.8.137.
84. Ji Y, Li H, Zhou ZK. *Paris caobangensis* YH Ji, H. Li & ZK Zhou (Trilliaceae), a new species from northern Vietnam. *Acta Phytotaxon Sin.* 2006;44:700–703. doi: 10.1360/aps050112.
85. Zachos J, Pagani M, Sloan L, Thomas E, Billups K. Trends, rhythms and aberrations in global climate 65 Ma to present. *Science.* 2001;292:686–693. doi: 10.1126/science.1059412.
86. Collinson ME. Vegetational and floristic changes around the Eocene/Oligocene boundary in western and central Europe. In: Prothero DR, Berggren WA, editors. *Eocene-Oligocene climate and biotic evolution.* Princeton: Princeton University Press; 1992. p. 437–450.
87. Santosh M. History of supercontinents and its relation to the origin of Japanese islands. *J Geodyn.* 2011;120:100–114. doi: 10.5026/jgeography.120.100.
88. Sun XJ, Wang PX. How old is the Asian monsoon system? Palaeobotanical records from China. *Palaeogeogr Palaeoclimatol Palaeoecol.* 2005;222:181–222. doi: 10.1016/j.palaeo.2005.03.005.
89. Wan SM, Li AC, Clift PD, Stut JBW. Development of the East Asian monsoon: mineralogical and sedimentologic records in the northern South China Sea since 20 Ma. *Palaeogeogr Palaeoclimatol Palaeoecol.* 2007;254:561–582. doi: 10.1016/j.palaeo.2007.07.009.
90. Jacques FMB, Guo SX, Su T, Xing YW, Huang YJ, Liu YS, et al. Quantitative reconstruction of the Late Miocene monsoon climates of southwest China: a case study of the Lincang flora from Yunnan Province. *Palaeogeogr Palaeoclimatol Palaeoecol.* 2011;304:318–327. doi: org/10.1016/j.palaeo.2010.04.014.

91. Zhang QQ, Ferguson DK, Mosbrugger V, Wang YF, Li CS. Vegetation and climatic changes of SW China in response to the uplift of Tibetan Plateau. *Palaeogeogr Palaeoclimatol Palaeoecol.* 2012;363:23–36. doi:10.1016/j.palaeo.2012.08.009.
92. Yao YF, Bruch AA, Mosbrugger V, Li CS. Quantitative reconstruction of Miocene climate patterns and evolution in southern China based on plant fossils. *Palaeogeogr Palaeoclimatol Palaeoecol.* 2011;304:291–307. doi: 10.1016/j.palaeo.2010.04.012.
93. Li FJ, Rousseau DD, Wu NQ, Hao QZ, Pei YP. Late Neogene evolution of the East Asian monsoon revealed by terrestrial mollusk record in western Chinese loess plateau: from winter to summer dominated subregime. *Earth Planet Sci Lett.* 2008;274:439–447. doi: 10.1016/j.epsl.2008.07.038.
94. An ZS, Kutzbach JE, Prell WL, Porter SC. Evolution of Asian monsoons and phased uplift of the Himalaya–Tibetan plateau since Late Miocene times. *Nature.* 2001;411:62–66. doi: 10.1038/35075035.
95. Yao T, Thompson L, Yang W, Yu WS, Gao Y, Guo XJ, et al. Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nat Clim Change.* 2012;2:663–667. doi:10.1038/nclimate1580.
96. Li XW, Li J. The Tanaka-Kaiyong line – an important floristic line for the study of the flora of East Asia. *Ann Mo Bot Gard.* 1997;84:888–892. doi: 10.2307/2992033.
97. Wang L, Schneider H, Zhang XC, Xiang QP. The rise of the Himalaya enforced the diversification of SE Asian ferns by altering the monsoon regimes. *BMC Plant Biol.* 2012;12:210. doi: org/10.1186/1471- 2229- 12-210.
98. Lu LM, Mao LF, Yang T, Ye JF, Liu B, Li HL, et al. Evolutionary history of the angiosperm flora of China. *Nature.* 2018;554:234–238. doi:10.1038/nature25485.
99. Harrison HC, Yu G, Takahara H, Prentice IC. Paleovegetation: diversity of temperate plants in East Asia. *Nature.* 2001;413:129–130. doi: 10.1038/35093166.
100. Royden LH, Burchfiel BC, Van DHRD. The geological evolution of the Tibetan Plateau. *Science.* 2008;321:1054–1058. doi: 10.1126/science.1155371.
101. Shi YF, Cui ZJ, Su Z. The Quaternary glaciations and environmental variations in China. Shijiazhuang: Hebei Science and Technology Publishing House; 2005.
102. Li JJ, Fang XM. Uplift of the Tibetan Plateau and environmental changes. *Chin Sci Bull.* 1999;44:2117–2124. doi: 10.1007/BF03182692.
103. Shi YF. Evolution of the cryosphere in the Tibetan Plateau, China, and its relationship with the global change in the mid-Quaternary. *J Glaciol Geocryol.* 1998; 20:197–208.
104. Axelrod DI, Al-Shehbaz I, Raven PH. History of the modern flora of China. In: Zhang AL, Wu SG, editors. Floristic characteristics and diversity of East Asian plants. Beijing: China Higher Education Press; 1998. p. 43–55.
105. Qiu YX, Fu CX, Comes HP. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Mol Phylogenet Evol.* 2011;59:225–244. doi: 10.1016/j.ympev.2011.01.012.

106. Qian H, Ricklefs RE. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature*. 2000;407:180–182. doi: 10.1038/35025052.
107. Wu ZY, Sun H, Zhou ZK, Peng H, Li DZ. Origin and differentiation of endemism in the flora of China. *Acta Bot Yunnan*. 2005;27:577–604. doi: 10.1007/s11515-007-0020-8.
108. Wen J, Zhang JQ, Nie ZL, Zhong Y, Sun H. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Front Genet*. 2014;5:4. doi: 10.3389/fgene.2014.00004.
109. Favre A, Päckert M, Pauls SU, Jähmig SC, Uhl D, Michalak I, et al. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biol Rev*. 2015;90:236–253. doi: 10.1111/brv.12107.
110. Vinogradov AE. Selfish DNA is maladaptive: evidence from the plant Red List. *Trends Genet*. 2003;19:609–614. doi: 10.1016/j.tig.2003.09.010.
111. Knight CA, Molinari NA, Petrov DA. The large genome constraint hypothesis: evolution, ecology and phenotype. *Ann Bot*. 2005;95:177–190. doi: 10.1093/aob/mci011.
112. Suda J, Kyncl T, Jarolímová V. Genome size variation in Macaronesian angiosperms: forty percent of the Canarian endemic flora completed. *Plant Syst Evol*. 2005;252:215–238. doi: 10.1007/s00606-004-0280-6.
113. Leitch IJ, Beaulieu JM, Chase MW, Leitch AR, Fay MF. Genome size dynamics and evolution in monocots. *Ann Bot*. 2010; doi: 10.1155/2010/862516.
114. Pellicer J, Kelly LJ, Leitch IJ, Zomlefer WB, Fay MF. A universe of dwarfs and giants: genome size and chromosome evolution in the monocot family Melanthiaceae. *New Phytol*. 2014;201:1484–1497. doi: 10.1111/nph.12617.
115. Valente LM, Savolainen V, Vargas P. Unparalleled rates of species diversification in Europe. *Proc R Soc B*. 2010;277:1489–1497. doi: 10.1098/rspb.2009.2163.
116. Klak C, Khunou A, Reeves G, Hedderson T. A phylogenetic hypothesis for the Aizoaceae (Caryophyllales) based on four plastid DNA regions. *Am J Bot*. 2003;90:1433–1445. doi: 10.2307/4123697.
117. Ji Y, Yang C, Huang Y. A new species of *Paris* sect. *Axiparis* (Melanthiaceae) from Yunnan, China. *Phytotaxa*. 2017;306:234–236. doi: 10.11646/phytotaxa.306.3.6.
118. Yin H, Zhang H, Xue D. *Paris polyphylla* var. *emeiensis* H.X.Yin, H.Zhang & D.Xue, a new variety of Trilliaceae from Sichuan, China. *Acta Phytotaxon Sin*. 2007;45:822–827. doi: 10.1360/aps06160.
119. Li H, Lei LG, Yang YM. *Paris yanchii*, a new species of *Paris* Linnaeus (Melanthaceae) from Yunnan. *J West China For Sci*. 2017;46:1–5. doi: 10.16473/j.cnki.xblykx1972.2017.01.001.
120. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 1987;19:11–15.
121. Patel RK, Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One*. 2012;7:e30619. doi: 10.1371/journal.pone.0030619.

122. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie2. *Nat Methods*. 2012;9:357–359. doi: 10.1038/nmeth.1923.
123. Wyman SK, Jansen RK, Boore JL. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 2004;20:3252–3255. doi: 10.1093/bioinformatics/bth352.
124. Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of *t*RNAs and *sno*RNAs. *Nucleic Acids Res*. 2005;33:W686–W689. doi: 10.1093/nar/gki366.
125. 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;28:1647–1649. doi: 10.1093/bioinformatics/bts199 .
126. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–780. doi: 10.1093/molbev/mst010.
127. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics*. 1998;14:817–818. doi: 10.1093/bioinformatics/14.9.817.
128. Posada D, Buckley TR. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Syst Biol*. 2004;53:793–808. doi:10.1080/10635150490522304.
129. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol*. 2017;34:772–773. doi: 10.1093/molbev/msw260.
130. Farris JS, Källersjö M, Kluge AC, Bult C. Testing significance of incongruence. *Cladistics*. 1994;10:315–319.
131. Swofford DL. PAUP: phylogenetic analysis using parsimony (and other methods), 4.0 beta. Massachusetts: Sinauer Associates; 2002.
132. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analysis with thousands of taxa and mixed models. *Bioinformatics*. 2006;22:2688–2690. doi: 10.1093/bioinformatics/btl446.
133. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003;19:1572–1574. doi: 10.1093/bioinformatics/btg180.
134. Givnish TJ, Zuluaga A, Marques I, Lam VKY, Gomez MS, Iles WJD, et al. Phylogenomics and historical biogeography of the monocot order Liliales: out of Australia and through Antarctica. *Cladistics*. 2016;32:581–605. doi: 10.1111/cla.12153.
135. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, et al. Beast 2: a software platform for bayesian evolutionary analysis. *PLoS Comput Biol*. 2014;10:e1003537. doi: 10.1371/journal.pcbi.1003537.
136. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 2019;35:526–528. doi: org/10.1093/bioinformatics/bty633.

137. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for statistical computing. Vienna: Austria; 2017.
138. Rabosky DL, Grudler M, Anderson C, Title P, Shi JJ, Brown JW, et al. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Meth Ecol Evol.* 2014;5:701–707. doi: 10.1111/2041-210X.12199.
139. Yu Y, Harris AJ, He XJ. S-DIVA (statistical dispersal-vicariance analysis): a tool for inferring biogeographic histories. *Mol Phylogenet Evol.* 2010;56:848–850. doi: 10.1016/j.ympev.2010.04.011.
- Yu Y, Harris AJ, Blair C, He X. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Mol Phylogenet Evol.* 2015;87:46–49. doi: 10.1016/j.ympev.2015.03.008.

## Table 1

**Table 1** Comparison of sequence characteristics of the aligned plastome and nrDNA datasets in *Paris*.

Dataset	Aligned length (bp)	Variable sites (divergence)	Parsimony informative sites (divergence)
Complete plastome	166,726	5,899 (3.53%)	3,225 (1.93%)
nrDNA	5,862	443 (7.56%)	264 (4.50%)

## Additional Files

**Additional file 1 (Table S1):** Summary of Illumina sequencing.

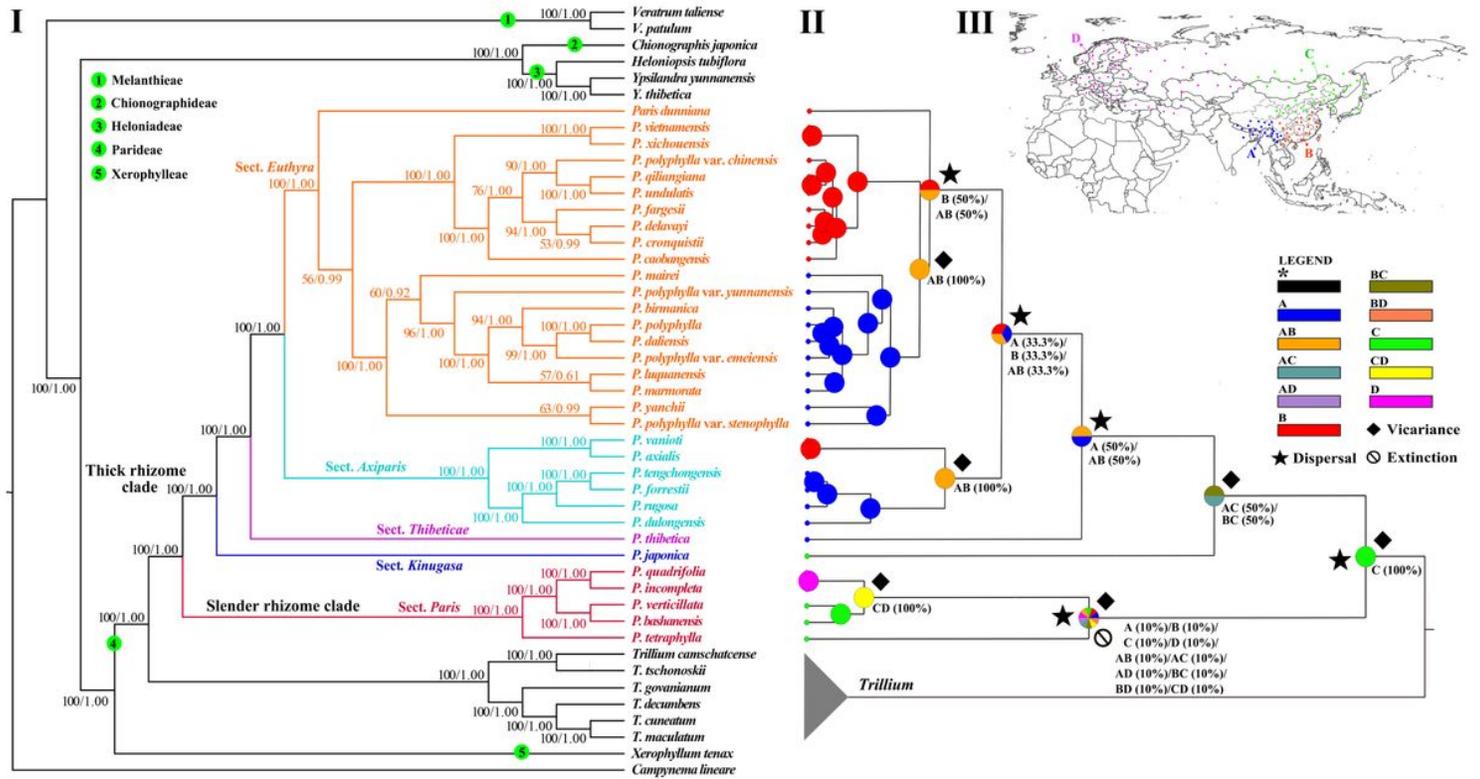
**Additional file 2 (Fig. S1):** Plastome map of *Paris* species.

**Additional file 3 (Table S2):** Gene content of *Paris* plastomes.

**Additional file 4 (Table S3):** Samples used in the study, with voucher and source information, and Genbank accessions.

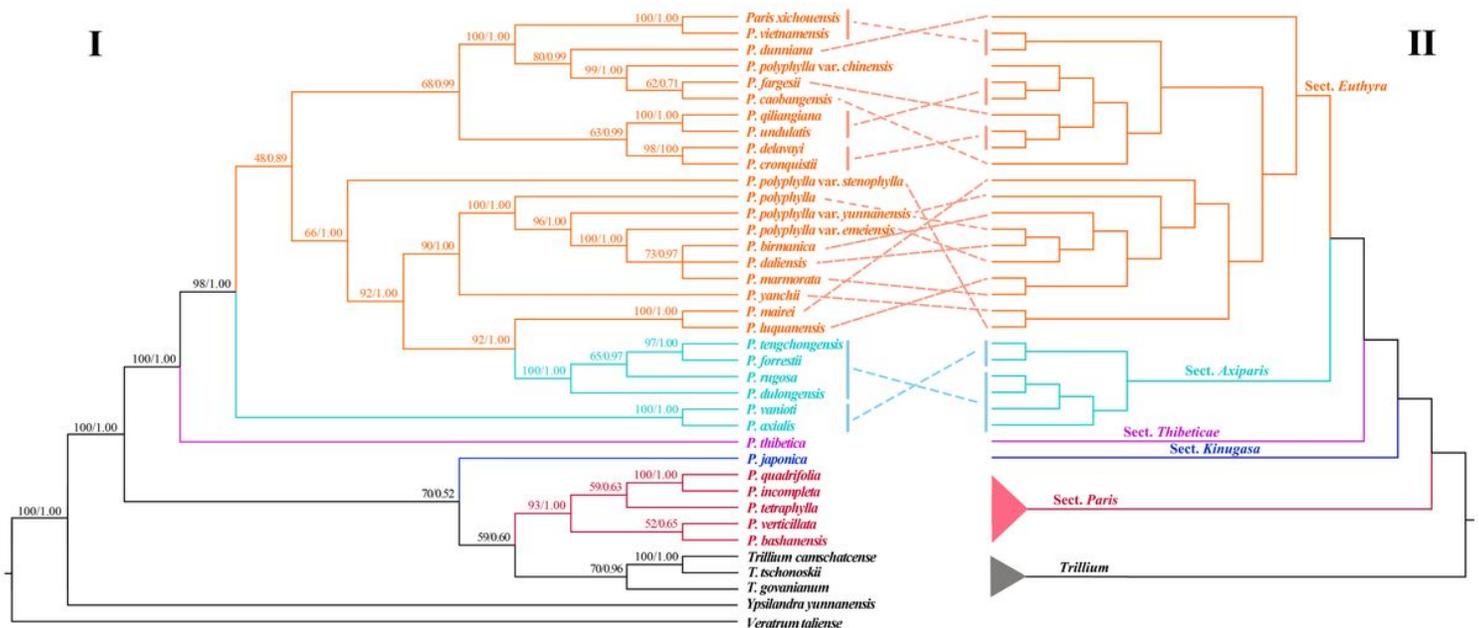
**Additional file 5 (Table S4):** Sequences downloaded from Genbank.

## Figures



**Figure 1**

Phylogenetic relationships and ancestral areas reconstruction of Paris. (I) Phylogenetic tree based on plastome DNA sequences. Numbers above branches indicate maximum likelihood bootstrap percentages (BP) and Bayesian posterior probabilities (PP). (II) Reconstruction of ancestral area of Paris using S-DIVA analysis inferred from plastid tree. (III) Paris species assigned to four areas based on their current distributions: A. southwestern China and Himalayas, B. eastern, central, southern China and northern Indochina, C. northeastern Asia and northern China, D. Europe and Caucasus.



**Figure 2**

Comparison of tree topologies recovered from analyses of nuclear ribosomal DNA sequences (I) and plastome DNA sequences (II). Numbers above branches indicated maximum likelihood bootstrap percentages (BP) and Bayesian posterior probabilities (PP).

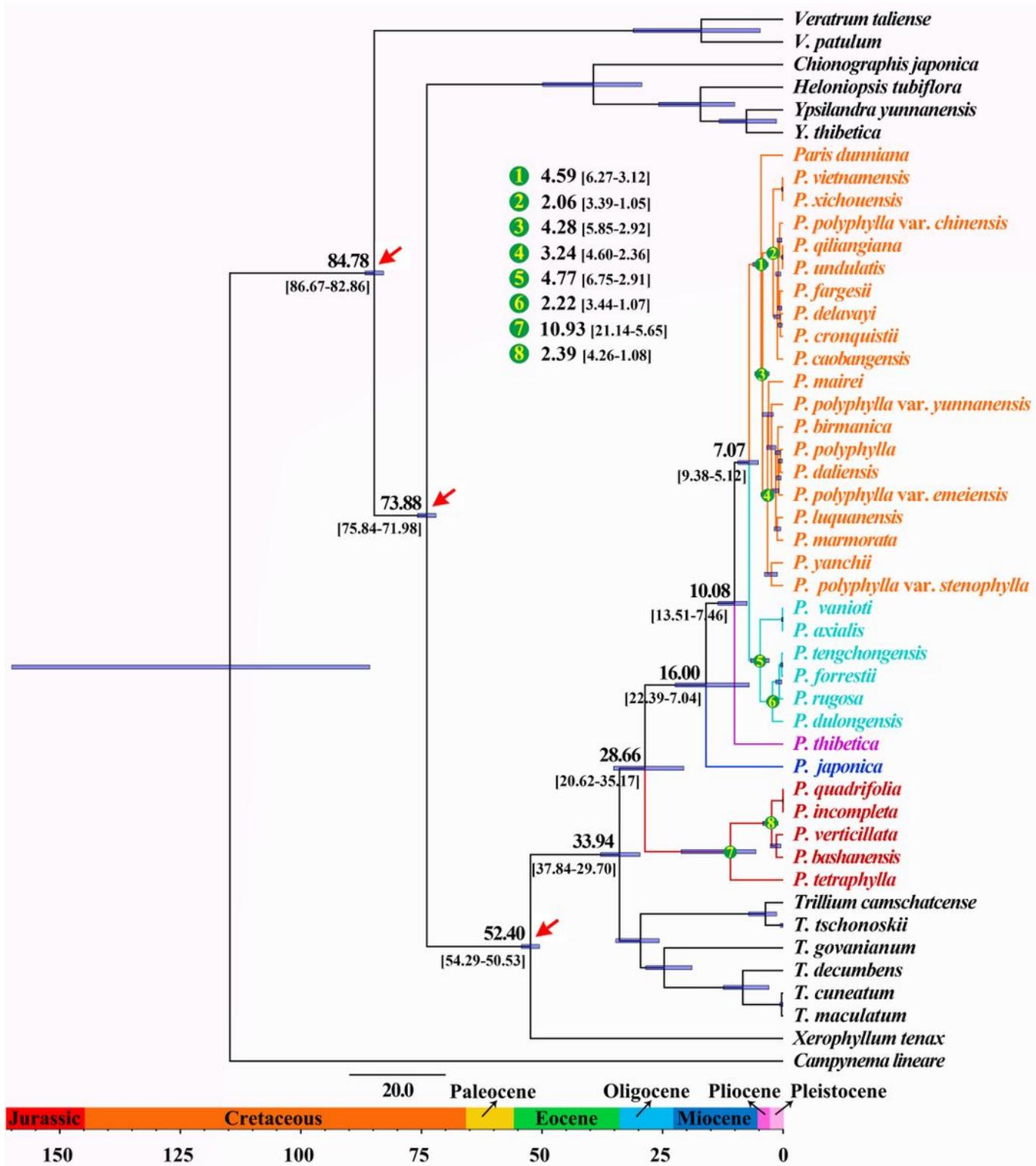
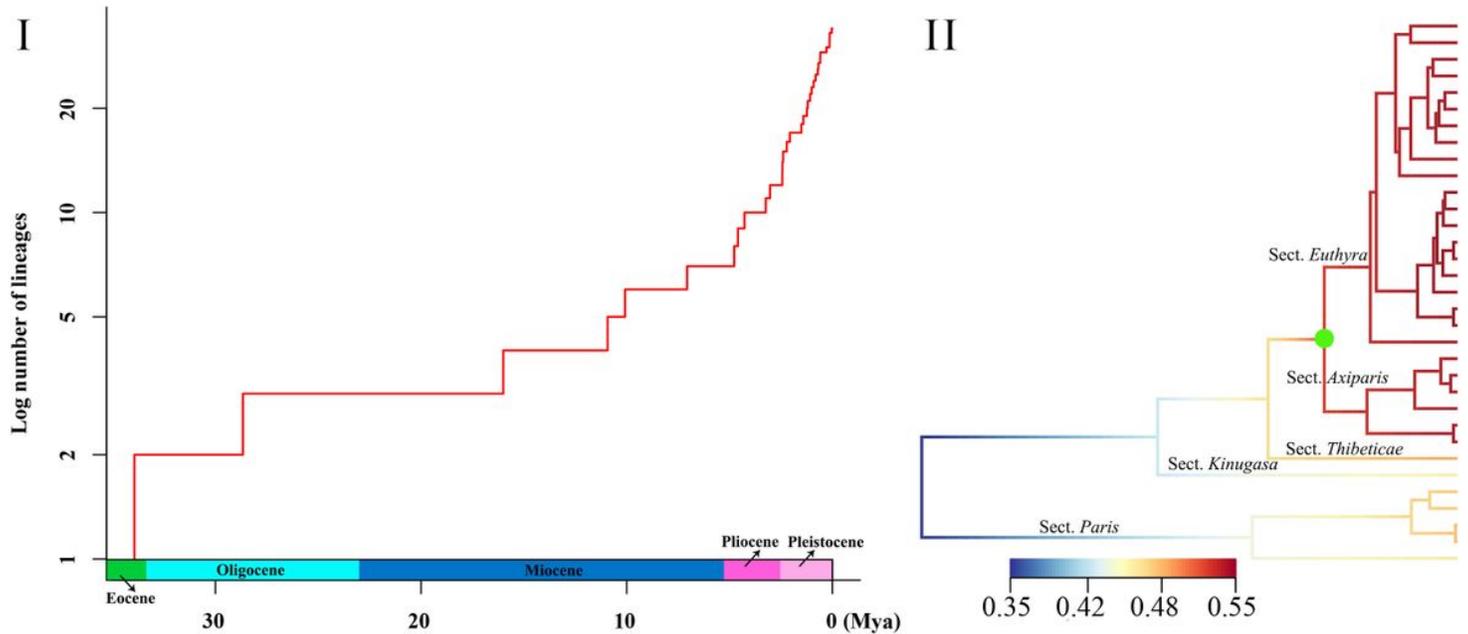


Figure 3

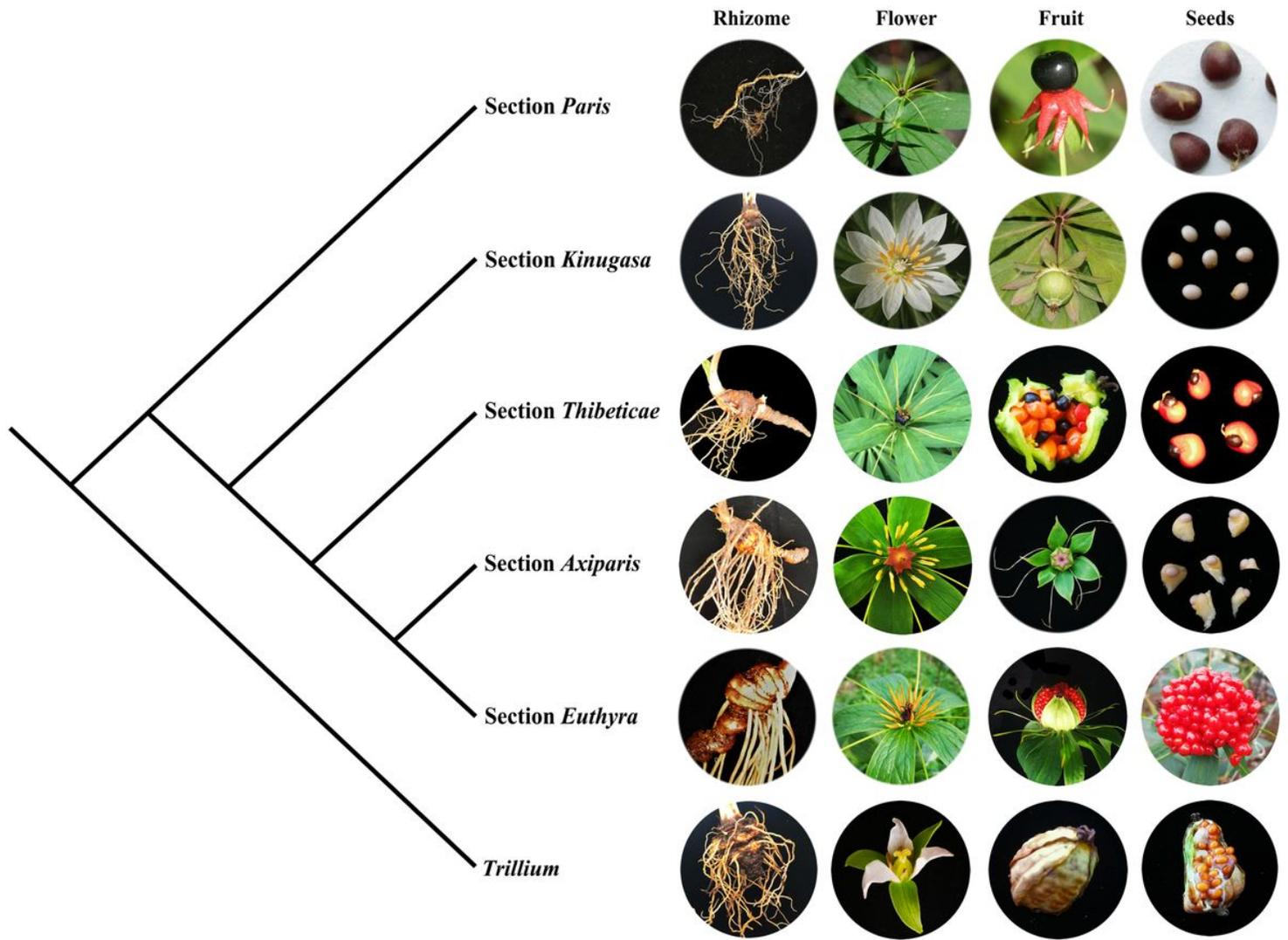
Divergence time estimation based on plastome DNA sequences. Numbers above/under the tree branches represented mean divergent ages and 95% confidence interval of each node. Red arrows indicate the

calibration points for the molecular dating. Divergence time and the timeline are indicated in million years (Mya).



**Figure 4**

(I) Result of lineage through time (LTT) plots (A) analyses. Mya = million years ago. (II) Rate shift with the highest posterior probability inferred in Bayesian Analysis of Macroevolutionary Mixtures (BAMM) on the time-calibrated maximum clade credibility tree from BEAST. A shift is predicted and shown with green circle. Colors on the branch represent the mean of the posterior density of net diversification rate (speciation rate minus extinction rate).



**Figure 5**

Comparison of morphological features among five Paris sections and Trillium.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS2submission.pdf](#)
- [FigureS1.jpg](#)
- [TableS1submission.pdf](#)
- [TableS4submission.pdf](#)
- [TableS3submission.pdf](#)