Complete Plastid Genome Sequencing of Trochodendraceae Reveals a Significant Expansion of the Inverted Repeat and Suggests a Paleogene Divergence between the Two Extant Species

Yan-xia Sun1,2, Michael J. Moore3, Ai-ping Meng1, Pamela S. Soltis4, Douglas E. Soltis4,5, Jian-qiang Li1*, Heng-chang Wang1*

1 Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, Hubei, China, 2 University of Chinese Academy of Sciences, Beijing, China, 3 Department of Biology, Oberlin College, Oberlin, Ohio, United States of America, 4 Florida Museum of Natural History, University of Florida, Gainesville, Florida, United States of America, 5 Department of Biology, University of Florida, Gainesville, Florida, United States of America

Abstract
The early-diverging euict order Trochodendraces contains only two monospecific genera, Tetracentron and Trochodendron. Although an extensive fossil record indicates that the clade is perhaps 100 million years old and was widespread throughout the Northern Hemisphere during the Paleogene and Neogene, the two extant genera are both narrowly distributed in eastern Asia. Recent phylogenetic analyses strongly support a clade of Trochodendraceae, Buxales, and Gunneridae (core euicots), but complete plastome analyses do not resolve the relationships among these groups with strong support. However, plastid phylogenomic analyses have not included data for Tetracentron. To better resolve basal euict relationships and to clarify when the two extant genera of Trochodendraceae diverged, we sequenced the complete plastid genome of Tetracentron sinense using Illumina technology. The Tetracentron and Trochodendron plastomes possess the typical gene content and arrangement that characterize most angiosperm plastid genomes, but both genomes have the same unusual ~4 kb expansion of the inverted repeat region to include five genes (rpl22, rps3, rpl16, rpl14, and rps8) that are normally found in the large single-copy region. Maximum likelihood analyses of an 83-gene, 88 taxon angiosperm data set yield an identical tree topology as previous plastid-based trees, and moderately support the sister relationship between Buxaceae and Gunneridae. Molecular dating analyses suggest that Tetracentron and Trochodendron diverged between 44-30 million years ago, which is congruent with the fossil record of Trochodendraceae and with previous estimates of the divergence time of these two taxa. We also characterize 154 simple sequence repeat loci from the Tetracentron sinense and Trochodendron aralioides plastomes that will be useful in future studies of population genetic structure for these relict species, both of which are of conservation concern.

Introduction
The euict order Trochodendraces [1] contains only two extant genera, both of which are monotypic: Trochodendron Sieb. & Zucc. and Tetracentron Oliver. Historically, these two genera have been treated either as the separate families Trochodendraceae and Tetracentraceae, or as the combined family Trochodendraceae [1–7]. The Trochodendraceae sensu APG III [1] appear to have been widespread in the Northern Hemisphere during the Paleogene and Neogene [7–15]. However, the two extant species of the family have small geographic ranges and are restricted to eastern Asia [16]. Trochodendron aralioides Sieb. & Zucc. is a large, evergreen shrub or small tree native to the mountains of Japan to South Korea and Taiwan, and the Ryukyu Islands [2,17], whereas Tetracentron sinense Oliver is a deciduous tree occurring in southwestern and central China and the eastern Himalayan regions. Both species are characterized by apetalous flowers arranged in cymose inflorescences and by loculicidal capsules that dehisce to release winged seeds [2,5,7,18]. Although earlier researchers reported that wood of Trochodendraceae wood lacked vessels and thus suggested that Trochodendraceae were among the earliest-diverging angiosperms, recent research has documented the presence of vessels in the wood of both genera [2,7,19].

Molecular phylogenetic studies, including analyses of complete plastid genome sequences, have routinely recovered Trochodendraces as an early-diverging member of the clade Euicots (sensa [20]; all italicized clad names follow this system), specifically as part of a strongly supported clade with Buxales and Gunneridae, or core euicots [21–27]. However, the relationships among Trochodendraces, Buxales, and Gunneridae have often been only
weakly supported. In the 17-gene analysis of Soltis et al. [28], which included data from all three plant genomes, Trochodendrales and Buxales were subsequent sisters to Gunneridae, with 100% and 98% BS support, respectively. However, other studies have found Buxales to be sister to Gunneridae with only weak support [24,26,29–30], whereas in other analyses Trochodendrales have appeared as sister to Gunneridae [27,31–32].

Complete plastid genome sequences have been used increasingly over the past decade to resolve deep-level phylogenetic relationships that have been unclear based on only a few genes. For example, recent plastid phylogenomic studies have helped to resolve key relationships among the earliest-diverging Mesangiospermae [33] as well as early-diverging Eudicotyledoneae and Pentapetalae [26,34]. Indeed, the plastid genome represents an excellent source of characters for plant phylogenetics due to the generally strong conservation of plastid genome structure and its mix of sequence regions that vary tremendously in evolutionary rate [35–37], which enable plastid genome sequence data to be applied to phylogenetic problems at almost any taxonomic level in plants [26,38–43]. It is now relatively inexpensive to generate complete plastid genome sequence data due to rapid improvements in next-generation sequencing (NGS) technologies [25,44–45] and due to the relatively small size of the plastid genome (∼150 kb) and its structural conservation, which enable dozens of plastomes to be multiplexed per sequencing lane and facilitate relatively straightforward genome assembly [45–48].

Figure 1. Map of the Tetracentron sinense plastid genome.
doi:10.1371/journal.pone.0060429.g001
Despite the promise of NGS technology for plastid genomics, the complete plastomes of only eight genera of early-diverging eudicots have been reported: *Ranunculus* (Ranunculaceae, Ranunculales), *Megaleranthis* (Ranunculaceae, Ranunculales), *Nandina* (Berberidaceae, Ranunculales), *Nelumbo* (Nelumbonaceae, Proteales), *Platanus* (Platanaceae, Proteales), *Meliosma* (Sabiaceae, Sabiales), *Trochodendron* (Trochodendraceae, Trochodendrales) and *Buxus* (Buxaceae, Buxales). Previous phylogenetic analyses based on some of these complete genomes have not fully resolved the relationships among early-diverging eudicots, however; in addition to the uncertainty surrounding relationships of Buxales, Trochodendrales, and *Gunneridae*, the positions of Sabiales and Proteales remain poorly supported [26–27]. Plastome taxon sampling is still sparse in these clades, however, and additional sampling may help elucidate these recalcitrant relationships.

**Figure 2. Map of the *Trochodendron aralioides* plastid genome.**

doi:10.1371/journal.pone.0060429.g002
In addition to their important role in phylogenetics, plastid genomes may be rich sources of population-level data. The non-recombination and uniparental inheritance of most plastid genomes can make plastid genomes extremely useful for population genetics, particularly for tracing maternal lineages [49–50]. For example, chloroplast simple sequence repeats (cpSSR) have been widely used in plant population genetics [51], including within early-diverging eudicots, where numerous cpSSR loci have been reported from the plastid genome of the endangered species Megaleranthis saniculifolia (Ranunculaceae) [52].

Here we report the complete plastid genome sequences of Tetracentron sinense and Trochodendron aralioides (the protein-coding and rRNA genes of Trochodendron cp genome were used for phylogenetic analyses in Moore et al. [26], but the cp genome structure of this genus has never been reported), as well as the results of new phylogenetic analyses based on adding Tetracentron and Megaleranthis genomes [52] to the 83-gene data set of Moore et al. [26]. We also compare the plastid genome structure of Trochodendron and Tetracentron, including the characterization of a significant expansion of the inverted repeat in both taxa, and we estimate the divergence time between the two genera. Finally, we characterize the distribution and location of cpSSRs in both Tetracentron sinense and Trochodendron aralioides, which provided further opportunity to study the population genetic structures of these two ancient relict species.

### Results

#### Sequencing and Genome Assembly

Illumina paired-end sequencing produced 892.11 Mb of data for Tetracentron sinense. We obtained 9912310 raw reads of 90 bp in length. The N50 of contigs was 13,981 bp and the summed length of contigs was 143,709 bp. The mean coverage of this genome was 5424.2×. After de novo and reference-guided assembly, we obtained a cp genome containing nine gaps. PCR and Sanger sequencing were used for filling the gaps. Four junction regions between IRs and SSC/LSC were first determined based on de novo contigs, and subsequently confirmed by PCR amplifications and Sanger sequencing, sequenced results were compared with the assembled genome directly and no mismatch or indel was observed, which validated the accuracy of our assembly. The genome sequences of Tetracentron sinense and Trochodendron aralioides have been submitted to GenBank (GenBank IDs: KC608752 and KC608753).

#### General Features of the Tetracentron and Trochodendron Plastomes

The plastid genome size of Tetracentron sinense is 164,467 base pairs (bp) (Figure 1), and that of Trochodendron aralioides is 165,945 bp (Figure 2). Both genomes show typical quadripartite structure, consisting of two copies of an inverted repeat (IR) separated by the large single-copy (LSC) and small single-copy (SSC) regions. The IRs contain the rps19 and rpl32 genes, which are transcribed in the opposite direction.

#### Table 2. The principal noncoding regions contributing to the size difference between the Tetracentron and Trochodendron plastid genomes.

| Spacer region or intron names | Tetracentron | Trochodendron | Length difference |
|------------------------------|--------------|---------------|------------------|
| trnK-UUU/rps16 spacer        | 870          | 1308          | 438              |
| rps16/trnQ-UUG spacer        | 1529         | 1797          | 268              |
| trnS-GCU/trnG-UCC spacer     | 505          | 658           | 153              |
| trnT-UGU/trnL-UAA spacer     | 957          | 1316          | 359              |
| petA/psbJ spacer             | 1146         | 754           | −392             |
| ycf1/ndhF spacer             | 440          | 325           | −115             |
| *rpl16 intron                | 865          | 972           | 107              |

All sizes are in base pairs. The only locus residing in the IR is marked with an asterisk (*). 

[doc:10.1371/journal.pone.0060429.t002]
Figure 4. Amount of sequence divergence between the protein-coding genes of *Tetracentron* and *Trochodendron*. doi:10.1371/journal.pone.0060429.g004

Figure 5. Sequence identity plot between *Trochodendron* and *Tetracentron*. doi:10.1371/journal.pone.0060429.g005
Characterization of SSR Loci

In all, 134 SSR loci (77 each from *Tetracentron sinense* and *Trochodendron aralioides*) were detected in the two plastid genomes, of which 123 are mononucleotide repeats, 26 are dinucleotide repeats, two are trinucleotide repeats, and one is a tetranucleotide repeat (Table 7). Nearly all of the SSR loci are composed of A/T repeats (Table 7), and these SSR loci are mostly present in noncoding regions. The tetranucleotide locus identified in *Tetracentron* is in the first intron of *ycf3*. The two trinucleotide loci in *Trochodendron* are both located in the spacer region between *trnK*-UGU and *rps16*. The unique C mononucleotide repeat from *Trochodendron* is present in the *trnF*-ndhC intergenic spacer region.

Phylogenetic and Molecular Dating Analyses

ML analyses of the 85-gene, 88-taxon data set yielded a tree with a similar topology and bootstrap support (BS) values (Figure 6) as that of the plastid phylogenomic study of Moore et al. [26]. The clades of *Trochodendron*+*Tetracentron* and *Ranunculus*+*Megaleranthis* were supported with 100% BS support. Trochodendrales are sister to the remaining angiosperms with high support (BS = 100%), but Buxaceae are sister to Gunneridae with only 67% BS support.

Molecular dating analyses suggest that *Trochodendron* and *Tetracentron* diverged between 44–30 million ago. The crown group 95% highest posterior density (HPD) age estimates for other major lineages of Pentapetalae were as follows: *Superasteridae* (115–109 mya), *Dilleniaceae+Cuperosidae* (116–112 mya), *Superrosidae* (114–111 mya), *Santalales* (98–75 mya), *Caryophyllales* (76–60 mya), *Asteridae* (104–99 mya), *Rosidae* (111–105 mya), *Vitaceae+Saxifragales* (114–110 mya), and *Saxifragales* (109–107 mya).
### Table 4. Comparisons of the protein-coding genes of *Tetracentron* and *Trochodendron*.

| Gene   | Length in *Tetracentron* | Length in *Trochodendron* | Number of nucleotide differences | Proportion of nucleotide differences | Number of indel differences |
|--------|--------------------------|----------------------------|----------------------------------|--------------------------------------|-----------------------------|
| petL   | 102                      | 102                        | 0                                | 0                                    | 0                           |
| psaI   | 111                      | 111                        | 0                                | 0                                    | 0                           |
| psaJ   | 129                      | 129                        | 0                                | 0                                    | 0                           |
| psbE   | 252                      | 252                        | 0                                | 0                                    | 0                           |
| psbF   | 120                      | 120                        | 0                                | 0                                    | 0                           |
| psbJ   | 123                      | 123                        | 0                                | 0                                    | 0                           |
| psbl   | 117                      | 117                        | 0                                | 0                                    | 0                           |
| psbT   | 108                      | 108                        | 0                                | 0                                    | 0                           |
| rpl23  | 288                      | 288                        | 0                                | 0                                    | 0                           |
| rps19  | 279                      | 279                        | 0                                | 0                                    | 0                           |
| rps7   | 468                      | 468                        | 0                                | 0                                    | 0                           |
| rps8   | 399                      | 399                        | 0                                | 0                                    | 0                           |
| rpl2   | 825                      | 825                        | 1                                | 0.00121                              | 0                           |
| rps3   | 657                      | 657                        | 1                                | 0.00152                              | 0                           |
| petD   | 504                      | 504                        | 1                                | 0.00198                              | 0                           |
| rpl16  | 501                      | 501                        | 1                                | 0.00249                              | 0                           |
| rpl14  | 369                      | 369                        | 1                                | 0.00271                              | 0                           |
| ycf2   | 6879                     | 6897                       | 19                               | 0.00276                              | 1                           |
| ndhB   | 1533                     | 1533                       | 5                                | 0.00326                              | 0                           |
| ycf3   | 507                      | 507                        | 2                                | 0.00394                              | 0                           |
| rpl33  | 201                      | 201                        | 1                                | 0.00498                              | 0                           |
| psbZ   | 189                      | 189                        | 1                                | 0.00529                              | 0                           |
| psaA   | 2253                     | 2253                       | 12                               | 0.00533                              | 0                           |
| psbK   | 186                      | 186                        | 1                                | 0.00538                              | 0                           |
| rps12  | 372                      | 372                        | 2                                | 0.00538                              | 0                           |
| psbA   | 1062                     | 1062                       | 6                                | 0.00565                              | 0                           |
| rpl20  | 354                      | 354                        | 2                                | 0.00565                              | 0                           |
| rpoC1  | 2049                     | 2070                       | 12                               | 0.00586                              | 1                           |
| atpA   | 1524                     | 1524                       | 9                                | 0.00591                              | 0                           |
| rpl22  | 486                      | 480                        | 3                                | 0.00625                              | 1                           |
| ndhJ   | 477                      | 477                        | 3                                | 0.00629                              | 0                           |
| psbD   | 1062                     | 1062                       | 7                                | 0.00659                              | 0                           |
| petA   | 963                      | 963                        | 7                                | 0.00727                              | 0                           |
| rpoB   | 3213                     | 3213                       | 24                               | 0.00747                              | 0                           |
| psbN   | 132                      | 132                        | 1                                | 0.00758                              | 0                           |
| psbB   | 2205                     | 2205                       | 17                               | 0.00771                              | 0                           |
| psbC   | 1422                     | 1422                       | 11                               | 0.00774                              | 0                           |
| atpH   | 246                      | 246                        | 2                                | 0.00813                              | 0                           |
| psaC   | 246                      | 246                        | 2                                | 0.00813                              | 0                           |
| ndhA   | 1095                     | 1095                       | 9                                | 0.00822                              | 0                           |
| rps4   | 606                      | 606                        | 5                                | 0.00825                              | 0                           |
| infA   | 234                      | 234                        | 2                                | 0.00855                              | 0                           |
| atpB   | 1497                     | 1497                       | 13                               | 0.00868                              | 0                           |
| cemA   | 690                      | 690                        | 6                                | 0.0087                                | 0                           |
| petG   | 114                      | 114                        | 1                                | 0.00877                              | 0                           |
| psbI   | 111                      | 111                        | 1                                | 0.00901                              | 0                           |
| ndhC   | 1428                     | 1428                       | 13                               | 0.00911                              | 0                           |
| petB   | 648                      | 648                        | 6                                | 0.00926                              | 0                           |
| atpI   | 744                      | 744                        | 7                                | 0.00941                              | 0                           |
Discussion

Expansion of the IR Region in Trochodendrales Plastomes

The plastid genomes of *Tetracentron* and *Trochodendron* exhibit the typical gene content and genome structure of angiosperms [37,53–54], with the notable exception of a significantly expanded IR region (Figures 1, 2, 3). This, a 4 kb expansion is responsible for the relatively large size of both Trochodendrales plastomes, which are ~4–5 kb larger than the typical upper size range of angiosperm plastid genomes, including those of nearly all other early-diverging eudicots (Table 8). Significant expansion, contraction, and even loss of the IR appears to be an evolutionarily uncommon phenomena but are nonetheless associated with much of the more significant variation in plastome size in angiosperms. For example, the largest known angiosperm plastome, that of *Pelargonium x hortorum*, also possesses the largest known IR, at ~76 kb in length [55]. Other significant IR expansions and contractions have been found in Campanulaceae [56–57], Apiaceae [58], and *Lemna* (Araceae) [59].

Impact of Additional Taxon Sampling on Basal Eudicot Phylogeny

The inclusion of *Megaleranthis* and *Tetracentron* in our analyses had no effect on the relationships among the major early-diverging eudicot lineages, and very little effect on support values. Of the basal splits among the eudicots with BS values less than 100% in both the current tree and that of Moore et al. [26], all were within 3% BS value. For example, the sister relationship of Buxales and Gunneridae is 70% in Moore et al. [26] vs. 67% with the inclusion of *Megaleranthis* and *Tetracentron*, and the sister relationship of Sabiales and Proteales has BS support of 80% in Moore et al. [26] vs. 83% in the current analyses. These similar values are unsurprising given that *Tetracentron* and *Trochodendron* are found to be relatively closely related in our analyses. Indeed, the relatively low sequence divergence between the *Tetracentron* and *Trochodendron* plastid genomes supports the taxonomic placement of Tetracentraceae within Trochodendrales, as advocated by APG III [1]. Although it is possible that the addition of the noncoding regions of the

| Gene | Length in *Tetracentron* | Length in *Trochodendron* | Number of nucleotide differences | Proportion of nucleotide differences | Number of indel differences |
|------|--------------------------|---------------------------|----------------------------------|-------------------------------------|-----------------------------|
| *clpP* | 609                      | 609                       | 6                                | 0.00985                             | 0                           |
| *rps14* | 303                      | 303                       | 3                                | 0.0099                             | 0                           |
| *atpE* | 402                      | 402                       | 4                                | 0.00995                            | 0                           |
| *ccsA* | 966                      | 966                       | 10                               | 0.01035                            | 0                           |
| *psbB* | 1527                     | 1527                      | 16                               | 0.01048                            | 0                           |
| *accD* | 1491                     | 1491                      | 16                               | 0.01073                            | 0                           |
| *ndhK* | 822                      | 858                       | 9                                | 0.01095                            | 1                           |
| *ndhC* | 363                      | 363                       | 4                                | 0.01102                            | 0                           |
| *petH* | 90                       | 90                        | 1                                | 0.01111                            | 0                           |
| *ndhG* | 531                      | 531                       | 6                                | 0.01122                            | 0                           |
| *rpoC2* | 4137                     | 4146                      | 50                               | 0.01209                            | 1                           |
| *ndhD* | 1503                     | 1503                      | 18                               | 0.01264                            | 0                           |
| *rps2* | 711                      | 711                       | 9                                | 0.01266                            | 0                           |
| *psbH* | 222                      | 222                       | 3                                | 0.01351                            | 0                           |
| *ndhI* | 543                      | 543                       | 8                                | 0.01473                            | 0                           |
| *atpF* | 555                      | 555                       | 9                                | 0.01622                            | 0                           |
| *matK* | 1536                     | 1536                      | 25                               | 0.01628                            | 0                           |
| *ndhE* | 306                      | 303                       | 5                                | 0.0165                             | 1                           |
| *rps18* | 303                      | 303                       | 5                                | 0.0165                             | 0                           |
| *ndhH* | 1182                     | 1182                      | 20                               | 0.01692                            | 0                           |
| *ycf4* | 555                      | 555                       | 10                               | 0.01805                            | 0                           |
| *rps15* | 273                      | 273                       | 5                                | 0.01832                            | 0                           |
| *psbM* | 105                      | 105                       | 2                                | 0.01905                            | 0                           |
| *rps11* | 417                      | 417                       | 9                                | 0.02158                            | 0                           |
| *rpoA* | 1014                     | 1014                      | 24                               | 0.02367                            | 0                           |
| *rpl32* | 162                      | 162                       | 4                                | 0.02469                            | 0                           |
| *rps16* | 227                      | 227                       | 6                                | 0.02622                            | 0                           |
| *ndhF* | 2223                     | 2223                      | 61                               | 0.02744                            | 0                           |
| *ycf1* | 5688                     | 5691                      | 195                              | 0.0345                             | 6                           |
| *rpl36* | 114                      | 114                       | 5                                | 0.04386                            | 0                           |

Genes are ranked from lowest to highest proportion of nucleotide differences. doi:10.1371/journal.pone.0060429.t004
plastid genome (or at least those noncoding regions that can be aligned) to our data set may improve support for these relationships, we may have to look to the other plant genomes for a confident resolution of relationships among the early-diverging eudicots. In fact, the sister relationship of Buxales and Gunneridae received high support (BS = 98%) in the 17-gene analyses of Soltis et al. [28], which employed a combination of 11 plastid genes, 18S and 26S nuclear rDNA, and 4 mitochondrial genes. However, the sister relationship of Sabiales and Proteales were more poorly supported (BS = 59%) in Soltis et al. [28].

Divergence Time Between Tetracentron and Trochodendron

Cenozoic Trochodendrales fossils are known throughout the Northern Hemisphere, with the Paleocene Nordenskiöldia the earliest certain fossil of the order [7–13]. Both Tetracentron and Trochodendron had wide distributions in the Northern Hemisphere during the Paleogene and Neogene. Fossil remains of Tetracentron have been found in Japan [60–61], Idaho [62], Princeton, British Columbia and Republic, Washington [63], and Iceland [15]; Trochodendron fossil remains have been reported from Kamchatka [64], Japan [11], Idaho and Oregon [11–12], Washington [7], and British Columbia [63]. Our estimate of the divergence time between the two genera of Trochodendrales (44-30 mya) encompasses the recent estimate of 37-31 mya from Bell et al. [65], which was based on analysis of 567 taxa and three genes, as well as the mid-Eocene estimate of 45 mya derived from the rbcL analysis of Anderson et al. [66], which employed numerous fossil constraints from the early-diverging eudicots. The congruence among these studies and with the fossil record suggests that a mid-to late Eocene divergence for the two extant Trochodendrales lineages may be a reasonable estimate.

Analysis of Plastid SSR Loci in the Trochodendrales

Because microsatellite loci, including cpSSRs, often exhibit high variation within species, they are considered valuable molecular markers for population genetics [67–69]. A limited number of SSR loci were recently characterized for Tetracentron [70], but no cpSSR loci are available for Trochodendrales. The 77 cpSSR loci that were identified in both Tetracentron and Trochodendron represent ~42% more loci than the 54 loci reported in the plastid genome of Megaleranthis (Ranunculaceae), the only other early-diverging eudicot for which a comprehensive analysis of cpSSR loci is available. The abundant and varied cpSSR loci identified in Trochodendrales will be useful in characterizing the population genetics of both extant species, which are of conservation interest in the wild because of their relatively narrow, presumably relictual distributions, and decreasing numbers [71]. Tetracentron is officially afforded second-class protection in China.

### Table 5. Exon and intron lengths (bp) in plastid genes containing introns in Tetracentron sinense and Trochodendron aralioides, respectively.

| Gene     | Exon 1 (Te/Tr) | Intron 1 (Te/Tr) | Exon 2 (Te/Tr) | Intron 2 (Te/Tr) | Exon 3 (Te/Tr) |
|----------|----------------|-----------------|----------------|-----------------|----------------|
| trnK-UUU | 37/37          | 35/35           |                |                 |                |
| trnG-UCC | 24/24          | 48/48           |                |                 |                |
| trnL-UAA | 35/35          | 50/50           |                |                 |                |
| trnV-UAC | 39/39          | 37/37           |                |                 |                |
| trnG-AUG | 42/42          | 35/35           |                |                 |                |
| trnA-UGC | 38/38          | 35/35           |                |                 |                |
| petB     | 6/6            | 642/642         |                |                 |                |
| petD     | 8/8            | 496/496         |                |                 |                |
| atpF     | 145/145        | 410/410         |                |                 |                |
| ndhA     | 553/553        | 542/542         |                |                 |                |
| ndhB     | 777/777        | 756/756         |                |                 |                |
| rpl2     | 391/391        | 434/434         |                |                 |                |
| rpl16    | 9/9            | 402/402         |                |                 |                |
| rps12    | 114/114        | 232/232         | 538/536        | 26/26           |
| rpoC1    | 432/432        | 1617/1638       |                |                 |                |
| ctp     | 71/71          | 292/292         | 659/650        | 246/246         |
| ycf3     | 124/124        | 230/230         | 731/758        | 153/153         |
| rps16    | 40/40          | 227/227         |                |                 |                |

The rps12 gene is trans-spliced, and hence the length of intron 1 is unknown. doi:10.1371/journal.pone.0060429.t005

### Table 6. A/T content (%) of different regions in Tetracentron and Trochodendron.

| Region   | Tetracentron | Trochodendron |
|----------|--------------|---------------|
| overall  | 61.86        | 61.98         |
| LSC      | 63.50        | 63.74         |
| IR       | 57.63        | 57.83         |
| SSC      | 67.84        | 67.48         |
| Protein-coding regions | 61.58 | 61.53 |

doi:10.1371/journal.pone.0060429.t006
Materials and Methods

Sample Preparation, Sequencing, and Assembly

Fresh leaves of *Tetracentron sinense* were collected from the Kunming Institute of Botany at the Chinese Academy of Sciences, and a voucher was deposited at the Herbarium of Wuhan Botanical Garden, Chinese Academy of Science (HIB). Chloroplast DNA was isolated following the protocol of Zhang et al. [45], and an Illumina library was constructed following the manufacturer’s protocol (Illumina). The DNA was indexed by tag and

| Base | Length | Position in plastid genome |
|------|--------|---------------------------|
| SSR loci in *Tetracentron* |
| A     | 10     | 2085–2094 17266–17275 118899–118910 162450–162461 163452–163463 163940–163951 |
|       | 11     | 9611–9621 47147–47157 50813–50823 75797–75807 80873–80883 82302–82312 133069–133079 160432–160442 |
|       | 12     | 217–228 49977–49988 50332–50343 162450–162461 163452–163463 163940–163951 |
|       | 14     | 65157–65170 |
|       | 15     | 38842–38856 |
|       | 17     | 39891–39907 |
|       | 18     | 74838–74855 |
|       | 22     | 72886–72907 |
| T     | 10     | 5266–5275 67706–67715 107277–107286 112508–112517 117373–117382 |
|       | 11     | 7004–7014 74779–74789 47779–47789 67810–67820 76013–76023 88492–88502 |
|       | 12     | 55307–55318 71732–71734 84983–84994 85471–85482 86473–86484 118884–118895 119027–119038 |
|       | 13     | 13902–13914 |
|       | 14     | 72926–72939 |
| AT    | 10     | 1734–1743 20842–20843 50404–50413 63181–63190 |
|       | 12     | 4862–4873 12996–13007 114822–114833 |
|       | 14     | 60686–60699 |
| TA    | 10     | 34083–34092 34111–34120 114741–114750 |
|       | 14     | 49132–49145 |
| TAAA  | 10     | 46875–46894 |
| SSR loci in *Trochodendron* |
| A     | 10     | 118854–118863 126258–126267 142993–143002 163821–163830 18142–18151 40389–40398 41065–41074 68969–68978 76681–76690 85629–85638 |
|       | 11     | 134406–134416 16427–16437 30306–30316 39963–39973 51490–51500 70911–70921 81823–81833 9789–9799 |
|       | 12     | 10420–10431 48058–48069 48322–48333 |
|       | 13     | 164932–164944 |
|       | 16     | 161805–161820 73777–73792 75726–75741 |
|       | 15     | 46189–46203 |
|       | 17     | 214–230 83299–83315 9304–9320 |
| T     | 10     | 108427–108436 120424–120433 121028–121037 122665–122674 131951–131960 164891–164900 20189–20198 40375–40387 48933–48942 53154–53163 53339–53348 5700–5709 6030–6039 68604–68613 72934–72943 83282–83291 87599–87608 |
|       | 11     | 127885–127895 14709–14719 55604–55614 57547–57557 |
|       | 12     | 50271–50282 |
|       | 13     | 73814–73826 86485–86497 |
|       | 14     | 76896–76909 |
|       | 15     | 48889–48903 |
|       | 16     | 89609–89624 |
| AT    | 10     | 1724–1733 51556–51565 64459–64468 |
|       | 12     | 4921–4932 4954–4958 4998–5009 5044–5055 5085–5096 5099–5110 5145–5156 5186–5197 5200–5211 |
|       | 18     | 73275–73292 |
| TA    | 10     | 1738–1747 21689–21698 |
| TAA   | 18     | 5016–5033 5218–5235 |
| C     | 10     | 55999–56008 |

Table 7. Distribution of SSR loci in the plastid genomes of *Tetracentron* and *Trochodendron*.
sequenced together with eight other species in one lane of an Illumina Genome Analyzer IIx at Beijing Genomics Institute (BGI) in Shenzhen, China. Illumina Pipeline 1.3.2 was used conducting image analysis and base calling. Raw sequence reads produced by Illumina paired-end sequencing were filtered for high quality reads which were subsequently assembled into contigs with a minimum length of 100 bp using SOAPdenovo [72] with the Kmer = 57. Contigs were aligned to the Trochodendron aralioides plastid genome using BLAST (http://blast.ncbi.nlm.nih.gov/), and aligned contigs were ordered according to the reference genome.

Genome Annotation and Analysis

The Tetracentron and Trochodendron plastid genomes were annotated with DOGMA [73] and BLAST tools from NCBI (the National Center for Biotechnology Information). Physical maps were generated using GenomeVx [74] with subsequent manual editing. Sequence divergence between the Tetracentron and Trochodendron plastid genomes was evaluated using DnaSP version 5.10 [75], and genome sequence identity plots were generated using mVISTA [76] (http://genome.lbl.gov/vista/mvista/submit.shtml). Msatfinder ver. 1.6.8 [77] was used to identify SSR loci by manually setting repeat units.

![Figure 6. A maximum likelihood tree determined by GARLI (−In L = −1095466.026) for the 83-gene, 88-taxon data set. Numbers associated with branches are ML bootstrap support values. Error bars around nodes correspond to 95% highest posterior distributions of divergence times based on 6 fossils using the program BEAST. Eo = Eocene, Mi = Miocene, Ol. = Oligocene, Pa = Paleocene, Pi = Pliocene. doi:10.1371/journal.pone.0060429.g006](image)

Table 8. Numbers of genes (including genes that span IR/SC junctions) in the IR regions of early-diverging eudicots.

| Basal eudicot lineages | Species                  | Genes in IR region | cp genome size (bp) |
|------------------------|--------------------------|--------------------|--------------------|
| Ranunculales           | Ranunculus macranthus    | 20                 | 155129             |
|                        | Megaleranthis saniculifolia | 19            | 159924             |
|                        | Nandina domestica        | 19                 | 156599             |
| Proteales              | Nelumbo lutea            | 18                 | 163206             |
|                        | Platanus occidentalis    | 19                 | 161791             |
| Sabiales               | Meliosma aff. cuneifolia | 18                 | 160357             |
| Buxales                | Buxus microphylla        | 18                 | 159010             |
| Trochodendrales        | Tetracentron sinense     | 24                 | 164467             |
|                        | Trochodendron aralioides | 24             | 165945             |

doi:10.1371/journal.pone.0060429.t008
Phylogenetic and Divergence Time Analyses

All protein-coding sequences, as well as all rRNA sequences, were extracted from the *Tetraenacon* and *Magalacanthus* plastome [52] and added manually to the 83-gene, 86-taxon alignment of Moore et al. [26]. ML analyses were performed on the concatenated 83-gene data set using the following partitioning strategy: (1) codon positions 1 and 2 together; (2) codon position 3; and (3) rRNA genes. The optimal nucleotide sequence model was selected for each partition using jModelTest 2.1.1 using the Decision Theory (DT) criterion [76]. The following models were selected: TVM+I+F for codon positions 1+2 and for codon position 3, and TIM1+I+F for rRNA.

Partitioned ML analyses were conducted using GARLI 2.0 [79]. A total of ten search replicates were conducted to find the optimal tree, and nonparametric bootstrap support was assessed with 100 replicates [80]. All ML searches used random taxon addition to build starting trees.

Divergence times were estimated using BEAST version 1.7.4 [81], using the same dating strategies employed in Moore et al. [26]. In addition to the three calibration points (used in Moore et al. [26]) of minimum ages of 131.8 mya for angiosperms [82–85], 125 mya for eudicots [83,86], and 85 mya for the most recent common ancestor of *Quercus* and *Camus* [26], we additionally constrained the stem lineage of Malpighiales using a minimum of 89.3 my [87] and the node uniting *Calycanthus* and *Liriodendron* using 90 my [88], and set the age of Protoeae to a minimum of 90 my [59].

Acknowledgments

We thank the anonymous reviewers for their helpful comments on earlier versions of this manuscript.

Author Contributions

Conceived and designed the experiments: JQJ, HJCW. Performed the experiments: YXS MJM APAM. Analyzed the data: YXS MJM. Contributed reagents/materials/analysis tools: YXS MJM JQL HCW. Wrote the paper: YXS MJM PSS DES HCW.

References

1. Angiosperm Phylogeny Group (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc 161: 105–121.
2. Smith AC (1945) A taxonomic review of *Trochodendron*. New York: Carnegie Museum of Natural History Press.
3. Cronquist A (1981) An Integrated System of Classification of Flowering Plants. New York: Columbia University Press.
4. Endress PK (1986) Reproductive structures and phylogenetic significance of extant primitive angiosperms. Plant Syst Evol 152: 1–28.
5. Fields PF (1996a) The Succor Creek flora of the middle Miocene Sucker Creek Formation, southwestern Idaho and eastern Oregon: Systematics and paleoecology. New York: Columbia University Press.
6. Magallon S, Crane PR, Herendeen PS (1999) Phylogenetic pattern, diversity, and diversification of eudicots. Ann Missouri Bot Gard 96: 297–372.
7. Pigg KB, Wehr WC, Iken-Bond SM (2001) *Tetraenacon* and *Nordenskioldia* (Trochodendraeaceae) from the Middle Eocene of Washington State, U.S.A. Int J Plant Sci 162: 1107–1136.
8. Crane PR (1989) Paleobotanical evidence on the early radiation of nonmagno- liid dicotyledons. Plant Syst Evol 162: 165–191.
9. Crane PR, Manchester SR, Dülcher DL (1990) A preliminary survey of fossil leaves and well-preserved reproductive structures from the Sentinel Butte Formation (Paleocene) near Almont, North Dakota. Fieldiana Geol NS 20: 1–65.
10. Crane PR, Manchester SR, Dülcher DL (1991) Reproductive and vegetative structure of *Nordenskioldia* (Trochodendraeaceae), a vesselless dicotyledon from the early Tertiary of the Northern Hemisphere. Ann J Bot 8: 1311–1334.
11. Manchester SR, Crane PR, Dülcher DL (1991) *Nordenskioldia* and *Trochodendron* fruits (Trochodendraeaceae) from the Miocene of northwestern North America. Bot Gaz 152: 357–368.
12. Fields PF (1996a) A Succor Creek flora of the middle Miocene Sucker Creek Formation, southwestern Idaho and eastern Oregon: Systematics and paleoecology. Ph.D. thesis, Michigan State University, East Lansing.
13. Fields PF (1996b) *A Trochodendron* infructescence from the 15 Ma Succor Creek flora in Oregon: a geographic and possibly temporal range extension. Ann J Bot 83: 105–110.
14. Manchester SR (1999) Biogeographical relationships of North American Tertiary Floras. Ann Missouri Bot Gard 86: 472–522.
15. Grimson F, Denk T, Zetter R (2008) Pollen, fruits, and leaves of *Tetraenacon* (Trochodendraeaceae) from the Cenomanian of Iceland and western North America and their palaeobiogeographical implications. Grana 47: 1–14.
16. Watson L, Dallwitz MJ (2006) The families of flowering plants: descriptions, illustrations, identification, information retrieval. Version 3.
17. Malbecher DJ (1987) The plant-book. Cambridge: Cambridge University Press.
18. Doweld AB (1996) Carpology, seed anatomy and taxonomic relationships of *Tetraenacon* (Tetraenaconaceae) and *Trochodendron* (Trochodendraeaceae). Ann J Bot 82: 413–443.
19. Li HF, Chaw SM, Du CM, Ren Y (2011) Vessel elements present in the secondary xylem of *Trochodendron* and *Tetraenacon* (Trochodendraeaceae). Flora 206: 595–600.
20. Caninio PD, Doyle JA, Graham SW, Judd WS, Olmstead RG et al. (2007) Towards a phylogenetic nomenclature of *Tetraenacon*. Taxon 56: 822–846.
21. Solis DE, Solis PS, Endress PK, Chase MW (2003) Phylogeny and Evolution of the Angiosperms. Sunderland, MA: Sinauer.
22. Qui Y, Dombrovski O, Lee J, Li L, Whithock BA, et al. (2005) Phylogenetic analysis of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. Int J Plant Sci 166: 815–842.
23. Qiu YL, Li L, Hendry TA, Li R, Taylor DW, et al. (2006) Reconstructing the basal angiosperm phylogeny: evaluating information content of mitochondrial genes. Taxon 55: 837–846.
24. Worberg A, Quandt D, Barniske A-M, Lohne C, Hihu KV, et al. (2007) Phylogeny of basal eudicots: insights from non-coding and rapidly evolving DNA. Org Divers Evol 7: 55–77.
25. Solís DE, Moore MJ, Burleigh JG, Bell CD, Solís PS (2010) Assembling the Angiosperm Tree of Life: progress and future prospects. Ann Missouri Bot Gard 97: 514–526.
26. Moore MJ, Solís PS, Bell CD, Burleigh JG, Solís DE (2010) Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc Natl Acad Sci USA 107: 4623–4628.
27. Moore MJ, Hassan N, Gitzendanner MA, Bruenm RA, Croley M, et al. (2011) Phylogenetic analysis of the plastid inverted repeat for 244 species: insights into deeper-level angiosperm relationships from a long, slowly evolving sequence region. Int J Plant Sci 172: 541–550.
28. Solís DE, Smith S, Cellinese N, Ruffalo-Rodriguez NF, Olmstead R, et al. (2011) Infering angiosperm phylogeny: a 17-gene analysis. Ann J Bot 98: 704–710.
29. Qiu YL, Li LB, Wang B, Chen Z, Knoopy V et al. (2006) The deepest diversities in land plants inferred from phylogenomic evidence. Proc Natl Acad Sci USA 103: 15311–15316.
30. Barniske AM, Boesch T, Muller K, Krug M, Worberg A, et al. (2012) Phylogenetics of early branching eudicots: Comparing phylogenetic signal across plastid introns, spacers, and genes. J Syst Evol 50: 85–100.
31. Hoot SB, Magallon S, Crane PR (1999) Phylogeny of basal eudicots based on three molecular data sets: atpB, rbcL and 18S nuclear ribosomal DNA sequences. Ann Mo Bot Gard 86: 1–32.
32. Solís DE, Solís PS, Chase MW, Most M, Albach D, et al. (2000) Angiosperm phylogeny inferred from a combined data set of 18S rDNA, rbcL, and atpB sequences. Bot J Linn Soc 130: 381–461.
33. Moore MJ, Bell CD, Solís PS, Solís DE (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc Natl Acad Sci USA 104: 19363–19368.
34. Jansen RK, Cai Z, Raubeson LA, Daniel H, DePamphilis CW, et al. (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA 104: 19369–19374.
35. Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84: 9034–9038.
36. Downie SR, Palmer JD (1992) Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Solís PS, Solís DE, eds. Molecular Systematics of Plants. New York: Chapman and Hall. 14–35.
37. Raubeson LA, Jansen RK (2005) Chloroplast genomes of plants. In: Henry R, ed. Diversity and Evolution of Plants-genotypic Variation in Higher Plants. Oxfordshire: CABl Publishing. 45–68.
38. Moore MJ, Dhingra A, Solís PS, Shaw R, Farmerie WG, et al. (2006) Rapid and accurate pyrosequencing of angiosperm plastid genomes. BMC Plant Biol 6: 17.
39. Parks M, Cronn R, Liston A (2009) Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol 7: 84.
41. Kumar S, Hahn FM, McManus CM, Cornish K, Whalen MC (2009) Comparative analysis of the complete sequence of the plastid genome of *Parthenium argentatum* and identification of DNA barcodes to differentiate *Parthenium* species and lines. BMC Plant Biol 9: 131.

42. Wu F-H, Chan M-T, Liao D-C, Hsu C-T, Lee Y-W, et al. (2010) Complete chloroplast genome of *Trachelium caeruleum* (Campanulaceae): multiple inversions, and several repeat families. Curr Genet 31: 419–429.

43. Steele PR, Hertweck KL, Mayfield D, McKain MR, Leebens-Mack J, et al. (2013) A targeted enrichment strategy for massively parallel sequencing of angiosperm plastid genomes. App Plant Sci: in press.

44. Zhang YJ, Ma PF, Li DZ (2011) High-Throughput Sequencing of Six Bamboo Chloroplast Genomes: Phylogenetic Implications for Temperate Woody Bamboos (Poaceae: Bambusoideae). PLoS ONE 6: e20596.

45. Stull GW, Moore MJ, Mandala VS, Douglas N, Kates H-R, et al. (2013) A molecular phylogeny of modern Silene (Caryophyllaceae). Am J Bot 83: 727–31.

46. Chelebaeva AI, Chigayeva GB (1988) The genus *Tetracentron* (Trochodendraceae), a Tertiary relict endemic to East Asia. Am J Bot 99: e320–e322.

47. Crane PR, Pedersen KR, Friis EM, Drinnan AN (1993) Early Cretaceous (early Campanian-Maastrichtian) leaves from the Potomac Group of North America. Syst Bot 18: 328–344.

48. Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, et al. (2012) Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. Am J Bot 99: 349–364.

49. McGuade DE, Stevens JE, Peroni PA, Ravelli JA (1996) The spatial distribution of chloroplast DNA and allozyme polymorphisms within a population of *Silene stenophylla* (Caryophyllaceae). Am J Bot 83: 727–31.

50. Small RL, Cronn RC, Wendel JF (2004) Use of nuclear genes for phylogeny reconstruction in plants. Aust Syst Bot 17: 145–70.

51. Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. Trends Ecol Evol 16: 142–147.

52. Steele PH, Hertweck KL, Mayfield D, McKain MR, Leebens-Mack J, et al. (2012) Quality and quantity of data recovered from massively parallel sequencing: Examples in Asparagus and Poaceae. Am J Bot 99: 330–348.

53. Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, et al. (2012) Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. Am J Bot 99: 349–364.

54. Palmer JD (1991) Plastid chromosomes: structure and evolution. In: Vasil IK. Molecular Biology of Plastids. San Diego, USA: Academic Press. 5–53.

55. Palm RJ, Bogorad L, eds. Cell Culture and Somatic Cell Genetics in Plants, Vol. 7A, The Genetics of Plants. London: Academic Press. 2004. 514 pages.

56. Hughes NF (1994) The Enigma of Angiosperm Origins. Cambridge, UK: Cambridge Univ Press.

57. Bremer GJ (1996) Flowering Plant Origin, Evolution and Phylogeny. New York: Chapman and Hall. 91–115.

58. Darriba D, Taboada GL, Doallo R, Posada D (2012) *jModelTest* 2: more models, new heuristics and parallel computing. Nat Methods 9: 772.

59. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

60. Friis EM, Pedersen KR, Crane PR (1999) Early angiosperm diversification: The diversity of pollen associated with angiosperm reproductive structures in Early Cretaceous floras from Portugal. Am Mo Bot Gard 46: 259–296.

61. Doyle JA (1992) Revised phylematic correlations of the lower Pteridaceae (USA) and the Cocobeach sequence of Gabon (Barremian-Aptian). Cretaceus 13: 337–349.

62. Hughes NF (1994) The Enigma of Angiosperm Origins. Cambridge, UK: Cambridge Univ Press.

63. Bremer GJ (1996) Flowering Plant Origin, Evolution and Phylogeny. New York: Chapman and Hall. 91–115.

64. Dooley KL, Leitch IJ, Morden CL, Towers H, Warter PE, et al. (2002) A phylogenetic analysis of angiosperm chloroplast genomes. Mol Biol Evol 19: 142–147.

65. Pigg KB, Dillsof RM, DeVore ML, Welc WC (2007) New diversity among the Trochodendraceae from the Early/Middle Eocene Okanogan Highlands of British Columbia, Canada, and northeastern Washington State, United States. Int J Plant Sci 168: 521–532.

66. Pigg KB, Dillsof RM, DeVore ML, Welc WC (2007) New diversity among the Trochodendraceae from the Early/Middle Eocene Okanogan Highlands of British Columbia, Canada, and northeastern Washington State, United States. Int J Plant Sci 168: 521–532.

67. Powell W, Morgante M, Medevitt R, Vendramini GG, Rafalski JA (1995) Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. Proc Natl Acad Sci USA 92: 7759–7763.

68. Grassi F, Lahra M, Science A, Imaizuo S (2002) Chloroplast SSR markers to assess DNA diversity in wild and cultivated grapevines. Vitis 41: 157–158.

69. Ebert D, Peakall R (2009) Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. Mol Ecol Res 9: 673–690.

70. Yang Z, Liu R, Tao C, Chen S, Ji Y (2012) Microsatellites for *Tetracentron sinense* (Trophodendraceae), a Tertiary relict endemic to East Asia. Am J Bot 99: e320–e322.

71. Fu LG (1992) China Plant Red Data Book (Vol.1). Beijing: Science Press.

72. Fu LG, Zhu H, Ruan J, Qian W, Fang X, et al. (2010) New de novo assembly of human genomes with massively parallel short read sequencing. Genome Res 20: 265–273.

73. Wyman SK, Janzen RK, Boone JL (2004) Automatic annotation of organelar genomes with DOGMA. Bioinformatics 20: 3252–3255.

74. Janzen RK, Boone JL (2004) Automatic annotation of organelar genomes with DOGMA. Bioinformatics 20: 3252–3255.

75. Rozas J, Sancho-Delbarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497.

76. Frazer KA, Pachter L, Polakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for comparative genomics. Nucleic Acids Res 32: W273–W279.

77. Thurston MI, Field D (2005) Masfinfinder: detection and characterisation of microsatellites, version 1.6.8.

78. Darriba D, Taboada GL, Doallo R, Posada D (2012) *jModelTest* 2: more models, new heuristics and parallel computing. Nat Methods 9: 772.

79. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

80. Berta RV, Pachter L, Leitch IJ, Towers H, Warter PE, et al. (2002) A phylogenetic analysis of angiosperm chloroplast genomes. Mol Biol Evol 19: 142–147.

81. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

82. Feuerman J (1965) Confidence limits on phylogeny: An approach using the bootstrap. Evolution 39: 783–791.

83. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29: 1196–1205.

84. Doyle JA (1992) Revised phylematic correlations of the lower Pteridaceae (USA) and the Cocobeach sequence of Gabon (Barremian-Aptian). Cretaceus 13: 337–349.

85. Hughes NF (1994) The Enigma of Angiosperm Origins. Cambridge, UK: Cambridge Univ Press.

86. Bremer GJ (1996) Flowering Plant Origin, Evolution and Phylogeny. New York: Chapman and Hall. 91–115.

87. Friis EM, Pedersen KR, Crane PR (1999) Early angiosperm diversification: The diversity of pollen associated with angiosperm reproductive structures in Early Cretaceous floras from Portugal. Am Mo Bot Gard 46: 259–296.

88. Doyle JA, Hutton GL (1991) Pollen and Spores: Patterns of Diversification. Oxford: Clarendon. 169–193.

89. Magallon S, Castillo A (2009) Angiosperm diversification through time. Am J Bot 96: 349–365.

90. Friis EM, Ekholm H, Pedersen KR, Crane PR (1999) Vigna/samhanthes spec. nov. – A calcicoleous flower from the Potomac Group [Early Cretaceous] of eastern North America. Int J Plant Sci 155: 772–785.

91. Crane PR, Pedersen KR, Friis EM, Drinnan AN (1993) Early Cretaceous (early to middle Albian) platanoid infructescences associated with *Aphelips* leaves from the Potomac Group of North America. Syst Bot 18: 320–344.