

Research Article

Complete Chloroplast Genome Sequence of Chinese Lacquer Tree (*Toxicodendron vernicifluum*, Anacardiaceae) and Its Phylogenetic Significance

Lu Wang ¹, Na He,² Yao Li,¹ Yanming Fang ¹ and Feilong Zhang²

¹Co-Innovation Center for Sustainable Forestry in Southern China, College of Biology and the Environment, Key Laboratory of State Forestry and Grassland Administration on Subtropical Forest Biodiversity Conservation, Nanjing Forestry University, Nanjing 210037, China

²Xi'an Raw Lacquer and Research Institute, Xi'an 710061, China

Correspondence should be addressed to Yanming Fang; jwu4@njfu.edu.cn

Received 7 September 2019; Revised 11 November 2019; Accepted 25 November 2019; Published 30 January 2020

Academic Editor: Fengjie Sun

Copyright © 2020 Lu Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chinese lacquer tree (*Toxicodendron vernicifluum*) is an important commercial arbor species widely cultivated in East Asia for producing highly durable lacquer. Here, we sequenced and analyzed the complete chloroplast (cp) genome of *T. vernicifluum* and reconstructed the phylogeny of Sapindales based on 52 cp genomes of six families. The plastome of *T. vernicifluum* is 159,571 bp in length, including a pair of inverted repeats (IRs) of 26,511 bp, separated by a large single-copy (LSC) region of 87,475 bp and a small single-copy (SSC) region of 19,074 bp. A total of 126 genes were identified, of which 81 are protein-coding genes, 37 are transfer RNA genes, and eight are ribosomal RNA genes. Forty-nine mononucleotide microsatellites, one dinucleotide microsatellite, two complex microsatellites, and 49 long repeats were determined. Structural differences such as inversion variation in LSC and gene loss in IR were detected across cp genomes of the six genera in Anacardiaceae. Phylogenetic analyses revealed that the genus *Toxicodendron* is closely related to *Pistacia* and *Rhus*. The phylogenetic relationships of the six families in Sapindales were well resolved. Overall, this study providing complete cp genome resources will be beneficial for determining potential molecular markers and evolutionary patterns of *T. vernicifluum* and its closely related species.

1. Introduction

Toxicodendron vernicifluum (Stokes) F. A. Barkley, commonly known as the Chinese lacquer tree, is a deciduous tree species with a toxic sap in the sumac family, Anacardiaceae [1]. The generic name of the species is derived from the Greek words *toxikos*, meaning “poison,” and *dendron*, meaning “tree,” while the specific name *vernicifluum* means “lacquer” in Latin [2]. *T. vernicifluum* is native to China and the Indian subcontinent and has been cultivated in other oriental countries, such as South Korea and Japan, for probably thousands of years [3–5]. Through tapping the trunk, the species can provide us with the raw lacquer, an excellent adhesive and painting material with multiple characteristics, such as anticorrosion, antirust, non-oxidation, acid resistance, alcohol resistance, and high-

temperature resistance [3, 6]. The lacquer is traditionally used to make various types of lacquerwares in China, Japan, South Korea, and several countries in Southeast and South Asia [1]. Furthermore, *T. vernicifluum* is sometimes used in Chinese medicine for the treatment of internal parasites and for stopping bleeding. Previous studies also reported that the urushiols of the species probably have an anticancer activity to human cancer cells, and the flavonoids extracted from its leaves possess therapeutic potentials as a multipotent agent against neurodegenerative diseases [7, 8].

Many recent studies of the Chinese lacquer tree have focused on secondary metabolites [9–11], anatomic features [12–14], growth traits [15–17], and classification and evaluation of cultivars of the species [18, 19]. Since 1978, more than 90 local cultivars, including some rare ones, have been recognized [20]. Those cultivars have different properties,

such as high yields, good quality of lacquer, and good timber [21]. However, controversy still surrounded the delimitation of the cultivars based on morphological and anatomic traits [22]. Although some researchers have used different types of molecular markers, such as nuclear microsatellites (nSSRs) and amplified fragment length polymorphism (AFLP), to describe the genetic variation of both natural populations and cultivars of *T. vernicifluum* [23, 24], the information on the genetic diversity and structure of the species was still limited by small sample size, narrow sampling range, and a few number of molecular markers. Thus, more studies on both nuclear and plastid genomes are expected to provide more useful markers to reveal the genetic variation pattern of *T. vernicifluum* and its cultivars in the future.

The genus *Toxicodendron* Mill. is well known for possessing urushiols, which can cause severe allergic contact dermatitis [25]. It consists of approximately 24 species with a disjunct distribution in temperate North America and eastern Asia [26]. Fifteen of them are native to China, mainly distributed in the regions south of the Yangtze River [27]. Previous studies have shown that *Toxicodendron* is a monophyletic group distinct from other genera of the *Rhus* complex [28]. Two temperate disjunct lineages have been recovered, one from section *Toxicodendron* and the other between the eastern North American *T. vernix* and the eastern Asian *T. vernicifluum*. The biogeographic history of the genus suggested that the Bering land bridge may have acted as the migration route that resulted in the current pattern of temperate disjunctions. Nonetheless, intrageneric relationships of *Toxicodendron* are still poorly understood [28]. Furthermore, Anacardiaceae is among the families of Sapindales, which is known for citrus, maples, lychees, mangos, and cashews [29]. Previous studies have shown that Sapindales is a monophyletic group [30, 31]. However, phylogenetic relationships among several families of the order, such as Simaroubaceae, Rutaceae, and Meliaceae, are still not fully resolved [30–32].

The chloroplast (cp) genome is nonrecombinant and uniparental and is more conserved than mitochondrial and nuclear genomes in terms of gene content, organization, and structure [33]. Also, the nucleotide substitution rate of chloroplast genes is higher than those of mitochondria genome but lower than those of nuclear genome [34]. The cp genome plays an important role in reconstruction of the green plant phylogeny and understanding the origins of economically important cultivated species and changes that have taken place during domestication [35, 36]. With the development of bioinformatics and high-throughput sequencing technology, various studies on the evolution of cp genomes have emerged in recent years [37].

Recently, the cp genome of *T. vernicifluum* cv. *Dahongpao*, a natural triploid cultivar, has been reported [38]. However, the differences of plastomes between diploid and triploid are rarely known. In this study, we sequenced and analyzed the complete cp genome of diploid *T. vernicifluum* and reconstructed the phylogeny of Sapindales based on 52 cp genomes of six families. The following questions were addressed: (1) What are the features of the cp genome of diploid *T. vernicifluum*? (2) How many potential

microsatellite markers can the cp genome provide for us? (3) What kinds of structural variation events have occurred across the cp genomes in Anacardiaceae? and (4) Can the cp genome information provide supporting data for phylogenetic reconstruction of both Anacardiaceae and Sapindales?

2. Materials and Methods

2.1. Sampling, DNA Extraction, and Illumina Sequencing. Healthy and fresh leaves of diploid *T. vernicifluum* were collected from an adult tree growing in the Lacquer Paint Research Institute, Shaanxi Province, China, during September 2018. The voucher specimen of the individual (voucher accession number LW20180905001) was stored at the Herbarium of Lacquer Paint Research Institute. Total genomic DNA was extracted using a modified CTAB method [39] and was fragmented by mechanical interruption (ultrasonic). After quality testing, the PCR-amplified library was sequenced with the Illumina HiSeq X Ten platform (Illumina Inc., San Diego, CA, United States) according to the manufacturer's manual.

2.2. Chloroplast Genome Assembling and Annotation. The NOVOplasty ver. 2.7.2 software [40] was used for the *de novo* assembly of the chloroplast genome of *T. vernicifluum* based on the cp-like reads extracted from a total of 25,291,737 high-quality sequences (Phred score ≥ 30) generated by Illumina sequencing. The CpGAVAS pipeline [41] was applied to annotate the protein-coding, rRNA, and tRNA genes. The tRNAscan-SE ver. 1.21 software [42] was used to verify the tRNA genes with default settings. The circular gene map was drawn by the OrganellarGenomeDRAW tool (OGDRAW) ver. 1.3.1 [43]. The relative synonymous codon usage (RSCU) value was estimated for each codon based on the coding sequences of 81 protein-coding genes using PhyloSuite ver. 1.1.152 [44]. Simple sequence repeats (SSRs) across the cp genome were determined by MISA (<http://pgrc.ipk-gatersleben.de/misa/misa.html>) [45], with the minimum number of repeats set to 10 for mononucleotide, to 6 for dinucleotide, and to 5 for tri-, tetra-, penta-, and hexanucleotide SSRs. We also used REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer/>) [46] to identify forward, reverse, complement, and palindromic repeats, with the minimum repeat size set to 8 and the hamming distance set to 1.

2.3. Genome Comparison. We used the publicly available chloroplast genome sequences of species from five genera of Anacardiaceae (last accessed 1 April 2019), i.e., *Rhus chinensis* Mill. (GenBank accession number KX447140), *Pistacia weinmannifolia* J. Poiss. ex Franch. (MF630953), *Anacardium occidentale* L. (KY635877), *Mangifera indica* L. (KY635882), and *Spondias mombin* L. (KY828469) to perform the comparative cp genomic analysis with *T. vernicifluum*. The sequence identity of those genomes was plotted using the mVISTA program [47] with LAGAN mode. Multiple genome alignments were conducted through the progressive Mauve algorithm [48] as implemented in the

Geneious software (Biomatters, Auckland, New Zealand) to detect the presence of large-scale evolutionary events such as rearrangement and inversion. The borders of large single-copy (LSC), small single-copy (SSC), and inverted repeat (IR) regions were visually displayed and compared among the six species using Irscope [49]. To detect the hotspots of intergeneric divergence, sequences of 100 common genes and 37 intergenic spacers were extracted for each species using Phylosuite ver. 1.1.152 [44] and aligned by MAFFT ver. 7.313 [50]. Following this step, the nucleotide diversity (P_i) value was calculated for each of the 137 loci using DnaSP ver. 6.12.03 [51].

2.4. Phylogenetic Analysis. We reconstructed the phylogeny of Sapindales based on 52 cp genomes representing 38 genera of six families (Table S1), including two genera of Burseraceae, two genera of Simaroubaceae, 10 genera of Meliaceae, six genera of Anacardiaceae, 10 genera of Sapindaceae, and nine genera of Rutaceae. Two species of Brassicales (*Carica papaya* L.) and Huerteales (*Tapiscia sinensis* Oliv.) (Table S1) were selected as outgroups following the interrelationships of orders recognized by APG IV [52]. These genomes cannot be aligned directly due to the occurrence of numerous differences regarding gene content and structure. Therefore, we used the HomBlocks pipeline [53] that automatically recognizes locally collinear blocks (LCBs) and excavates phylogeny informative regions to construct the multigene involved alignment. In brief, the progressive Mauve algorithm [48, 54] was applied at first to identify coexisting blocks. Those LCBs were extracted and then trimmed by Gblocks [55]. The circoletto webserver (<http://tools.bat.infospire.org/circoletto/>) was used for the visualization of genes that were integrated into the final alignment. The best-fit partitioning schemes and DNA substitution models were determined by PartitionFinder [56] using a greedy search strategy. Finally, the GTR+G model was chosen for the subset including LCBs 2-3, and the TVM+I+G model was chosen for the subset including LCBs 4-6, the subset including LCBs 8-9, and the subset including only LCB5 (Table S2). The phylogenetic relationships among those 54 species were reconstructed using Bayesian-inference (BI) analyses as implemented in MrBayes ver. 3.2.6 [57]. A Markov chain Monte Carlo (MCMC) was run for 2,000,000 generations with two parallel searches using four chains, each starting with a random tree. Trees were sampled every 100 generations and the first 25% was discarded as burn-in. The maximum likelihood (ML) analyses were also performed with RaxmlGUI ver. 1.5b1 [58, 59]. The GTRGAMMA model was selected for all the subsets, and branch support values were estimated for each node based on 1,000 samples for rapid bootstrap.

3. Results

3.1. Features of *T. vernicifluum* Chloroplast Genome. A total of 25,291,737 paired-end reads were produced by the Illumina Hiseq X Ten sequencing platform. After *de novo* assembly, the complete cp genome sequence of *T. vernicifluum*

was obtained and submitted to the NCBI database under the GenBank accession number MK419151. The cp genome of *T. vernicifluum* is a circular molecule of 159,571 bp, consisting of a large single-copy (LSC) region of 87,475 bp, a small single-copy (SSC) region of 19,074 bp, and a pair of inverted repeats (IRa and IRb) of 26,511 bp (Table 1 and Figure 1). The base composition of the complete cp genome sequence was analyzed and found to be 31.30% T, 30.74% A, 19.06% C, and 18.90% G. The overall GC content was 37.96%, which is very close to those of other Anacardiaceae species, e.g., *M. indica* (37.89%), *R. chinensis* (37.79%), and *A. occidentale* (38.12%) (Table 1) and also to those of other species in Sapindales, e.g., *Toona ciliata* M. Roem. (37.90%) [60], *Commiphora gileadensis* (L.) C. Chr. (37.90%) [61], and *Xanthoceras sorbifolium* Bunge (37.70%) [62]. Furthermore, the GC contents are unevenly distributed across regions of the cp genome, which were found to be 36.08%, 32.63%, and 42.96% for the LSC, SSC, and IR regions, respectively (Table 2).

The complete cp genome of *T. vernicifluum* encodes 126 predicted functional genes, including 81 protein-coding genes, 37 transfer RNA (tRNA) genes, and eight ribosomal RNA (rRNA) genes (Table S3). Among those, 107 are unique, and 19 are duplicated in the IR regions. For the 107 unique genes, 30 are tRNA genes, four are rRNA genes, and 73 are protein-coding genes. For the 19 duplicated genes, eight are protein-coding genes, seven are tRNA genes, and four are rRNA genes. Within the cp genome of *T. vernicifluum*, 12 genes, including six tRNA genes and six protein-coding genes, contain only one intron. Two genes (*ycf3* and *clpP*) contain two introns (Table S4).

3.2. Codon Usage Bias. The relative synonymous codon usage (RSCU) value was estimated for each codon based on the coding sequences (CDS) of 81 protein-coding genes, which presented a total length of 80,280 bp and accounted for 50.31% of the complete cp genome of *T. vernicifluum*. A total of 26,760 codons were found in those coding regions. The most prevalent amino acid is leucine (2,794 codons, approximately 10.44%), followed by isoleucine (2,249 codons, approximately 8.40%) and serine (2,107 codons, approximately 7.87%), while the rarest one is cysteine (318 codons, approximately 1.19%) (Table S5). In addition, almost all of the A/U-ending codons showed RSCU values greater than 1, whereas the C/G-ending codons showed RSCU values less than 1. For example, synonymous codons GUU, GUC, GUA, and GUG encode valine, and the corresponding RSCU values for these four codons are 1.43, 0.51, 1.49, and 0.57, respectively, as expected from the low GC content of CDS.

3.3. Repeat Sequences. Simple sequence repeats (SSRs) are DNA stretches consisting of short, tandemly repeated motifs with a length of 1–6 bp, which have been widely used as molecular markers in population genetics and evolutionary biology [63–65]. In this study, a total of 52 chloroplast SSRs (cpSSRs) were identified ranging in length from 10 to 78 bp. Among those, 44 are distributed in intergenic spacers (IGSs),

TABLE 1: Summary of features of six Anacardiaceae chloroplast genomes.

Genome features	<i>Toxicodendron vernicifluum</i>	<i>Rhus chinensis</i>	<i>Pistacia weinmannifolia</i>	<i>Anacardium occidentale</i>	<i>Mangifera indica</i>	<i>Spondias mombin</i>
Cp length	159,571	149,011	160,767	172,199	157,780	162,302
LSC length	87,475	96,882	88,402	87,727	86,673	89,938
SSC length	19,074	18,647	19,129	19,046	18,349	18,012
IR length	26,511	16,741	26,618	32,713	26,379	27,176
<i>ycf2-trnL</i> length	1,010	1,083	1,011	7,251	1,011	1,018
Genes	126	126	132	129	128	130
CDS	81	82	87	84	83	86
tRNA	37	36	37	37	37	36
rRNA	8	8	8	8	8	8
GC%	37.96	37.79	37.84	38.12	37.89	37.60

six are located at coding regions of three genes (*rpoC2*, *atpB*, and *ycf1*), and only two are within the introns of *ycf3* and *clpP* (Table S6). On the other hand, most SSRs are distributed in LSC and SSC, whereas only 11 are located at IR regions (Table S6).

Mononucleotide SSRs were found to be the richest, with 49 of 52 belonging to this type. One of the other three is dinucleotide microsatellite with (AT)_n repeats, and two are complex microsatellites that include a small insertion and/or contain two different types of repeats (Table S6). No tri-, tetra-, penta-, and hexanucleotide SSRs were detected. Among the mononucleotide SSRs, polyadenine (poly A) and polythymine (poly T) occupy the highest portion (91.84%) and mononucleotide C and G are rare (8.16%). The higher level of A or T within those cpSSRs is consistent with the overall A/T content (62.04%) of the cp genome of *T. vernicifluum*. These SSR loci could be used for investigations into genetic diversity and genetic structure of natural populations and cultivars of *T. vernicifluum*.

Long repeats may promote the cp genome rearrangement and increase the population genetic diversity [66]. Using REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>), 49 long repeats were identified across the *T. vernicifluum* plastome, ranging from 22 to 66 bp in length (Table S7). Among those, 23 are forward repeats, 22 are palindromic repeats, and four are reverse repeats. Multiple nested sequence repeats were detected in *ycf2*.

3.4. Genome Comparison of Anacardiaceae Species. To investigate the intergeneric divergence of cp genome sequences, the percentage of identity was plotted for six species from different genera in Anacardiaceae using the mVISTA program with *T. vernicifluum* as the reference. High similarity was detected among those six genera, and the IR regions were found to be more conserved than the LSC/SSC regions (Figure 2). Furthermore, the variation of noncoding regions is significantly higher than that of coding regions (Figure 2).

The nucleotide diversity (P_i) values were calculated for 137 loci (100 genes and 37 intergenic spacers) to determine the hotspots of divergence. Those values ranged from 0 to 0.154, with a higher level of genetic variation detected within the LSC (average $P_i = 0.025$) and SSC regions (average

$P_i = 0.037$) than within IR regions (average $P_i = 0.003$) (Figure 3). In addition, intergenic spacers (average $P_i = 0.053$) were found to be more variable than genes (average $P_i = 0.013$) (Figure 3). Six of these loci, including *trnH-psbA* (0.154), *trnT-trnL* (0.154), *atpF-atpH* (0.126), *ccsA-ndhD* (0.096), *petD-rpoA* (0.083), and *trnL-trnF* (0.078), showed high levels of nucleotide diversity ($P_i > 0.070$) across the six genera of Anacardiaceae (Figure 3).

To characterize the structure and synteny, the cp genome sequences of the six Anacardiaceae species were aligned by Mauve, and *T. vernicifluum* was used as a reference to compare the gene orders among these cp genomes (Figure 4). Synteny analysis indicated that no rearrangement events were identified. However, differences were still found in terms of inversion variation in LSC and gene loss in IRs, leading to the occurrence of four collinearity modules across the six cp genome sequences. First, a large inverted segment was detected for *M. indica*, which resulted in the separation of the first block from the third block. This inversion was between *trnE-UUC* and *trnL-UAA*, with a length of 16,910 bp (ranging from 33,130 to 50,039 bp). Fourteen genes were included in this fragment, and two tRNA genes (*trnT-GGU* and *trnT-UGU*) were observed at the two ends. Second, due to the loss of a large fragment (~9,800 bp) in the IRb region of *R. chinensis*, the remaining part of the cp genome was divided into two blocks (the third and the fourth blocks). Instead, if the plastome of *R. chinensis* was removed, only three collinearity modules were identified across the cp genome sequences of the rest five species (Figure S1). This event has reduced the IR length of *R. chinensis* severely (Table 1) and shifted four genes (*ycf2*, *trnI-CAU*, *rpl23*, and *rpl2*) from IR to LSC regions.

3.5. IR Expansion and Contraction. For the six Anacardiaceae species, the length of IR was found to be significantly correlated with the total length of the complete cp genome ($R = 0.962$, $P = 0.002$) (Figure S2). Among those, *R. chinensis* presented the shortest length of cp genome, which was mainly attributed to the loss of a long fragment (~9,800 bp in length) at the IRb/LSC boundary. By contrast, *A. occidentale* showed the longest length of cp genome, which was associated with the size expansion of *ycf2-trnL* spacer (~6,200 bp

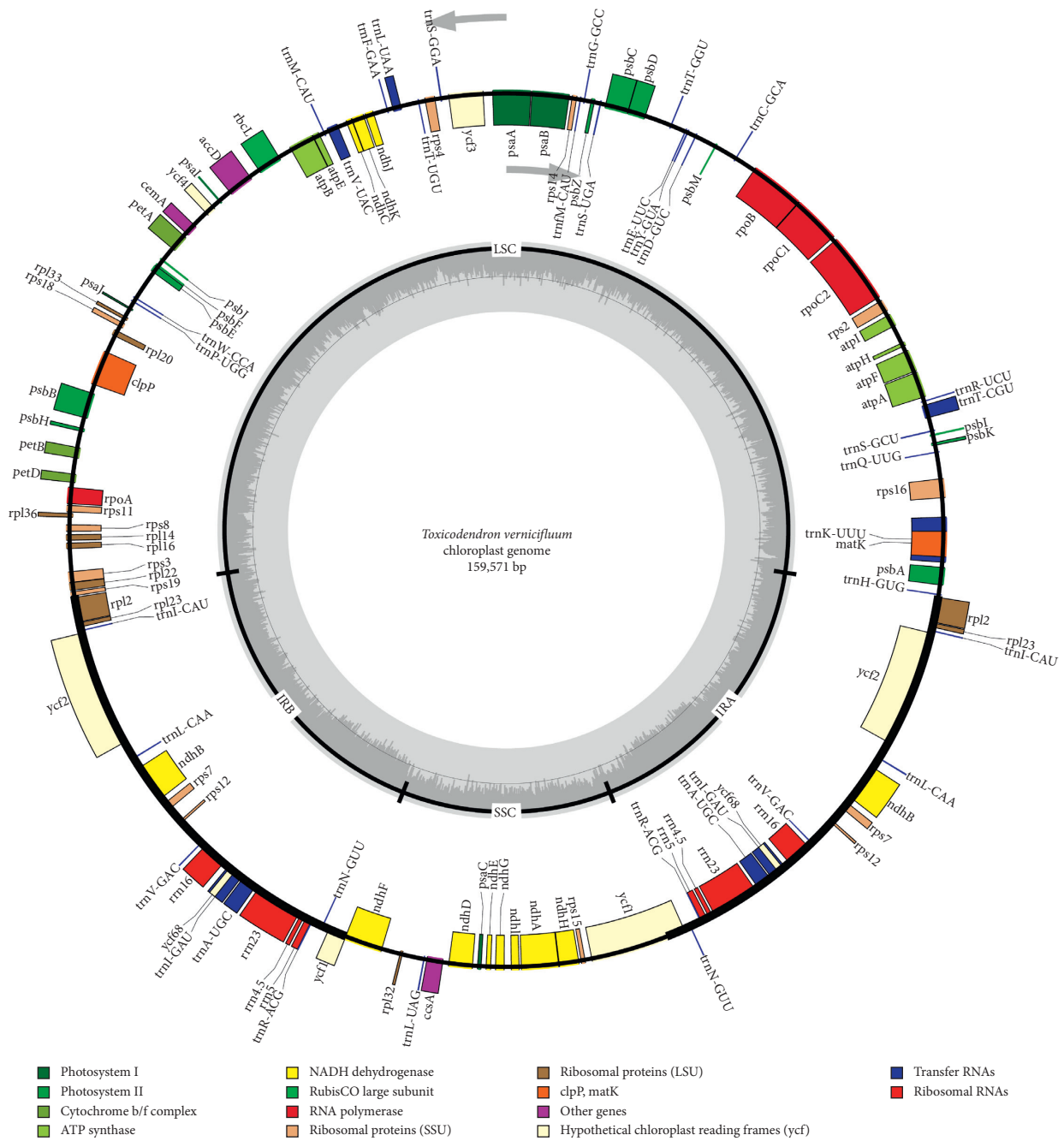


FIGURE 1: Gene map of *Toxicodendron vernicifluum* chloroplast genome. Genes shown inside the circle are transcribed clockwise and those outside are transcribed counterclockwise. Genes of different functions are color-coded. The darker gray in the inner circle shows the GC content, while the lighter gray shows the AT content.

TABLE 2: Base composition of the *Toxicodendron vernicifluum* chloroplast genome.

Region	A (%)	T (U) (%)	C (%)	G (%)	AT (%)	GC (%)
LSC	31.33	32.59	18.53	17.55	63.92	36.08
SSC	33.83	33.54	16.97	15.67	67.37	32.63
IR	28.51	28.51	21.48	21.48	57.04	42.96
Total	30.74	31.31	19.06	18.90	62.04	37.96

longer than the other five species) within the IR regions (Table 1). Furthermore, minor shifts of IR/SC boundaries were detected for those species (Figure 5). For example, The

rps19 gene is across the IRb/LSC border with 2–275 bp extending into the LSC region, while it was found to be 103 bp away from the border of *P. weinmannifolia*. Similarly, the *trnH* gene is located in the LSC region, 38–162 bp away from the IRa/LSC border, while it extends only 1 bp into the IRa region of *P. weinmannifolia*. For all the six species, the IRa/SSC junction is located in the *ycf1* region, and the IRb/SSC junction is located in the *ndhF* region. However, the length included in the IRs varies significantly across species, namely, 1,095–1,440 bp for *ycf1* and 12–42 bp for *ndhF*. Previous studies have shown that pseudogene *ycf1* overlaps

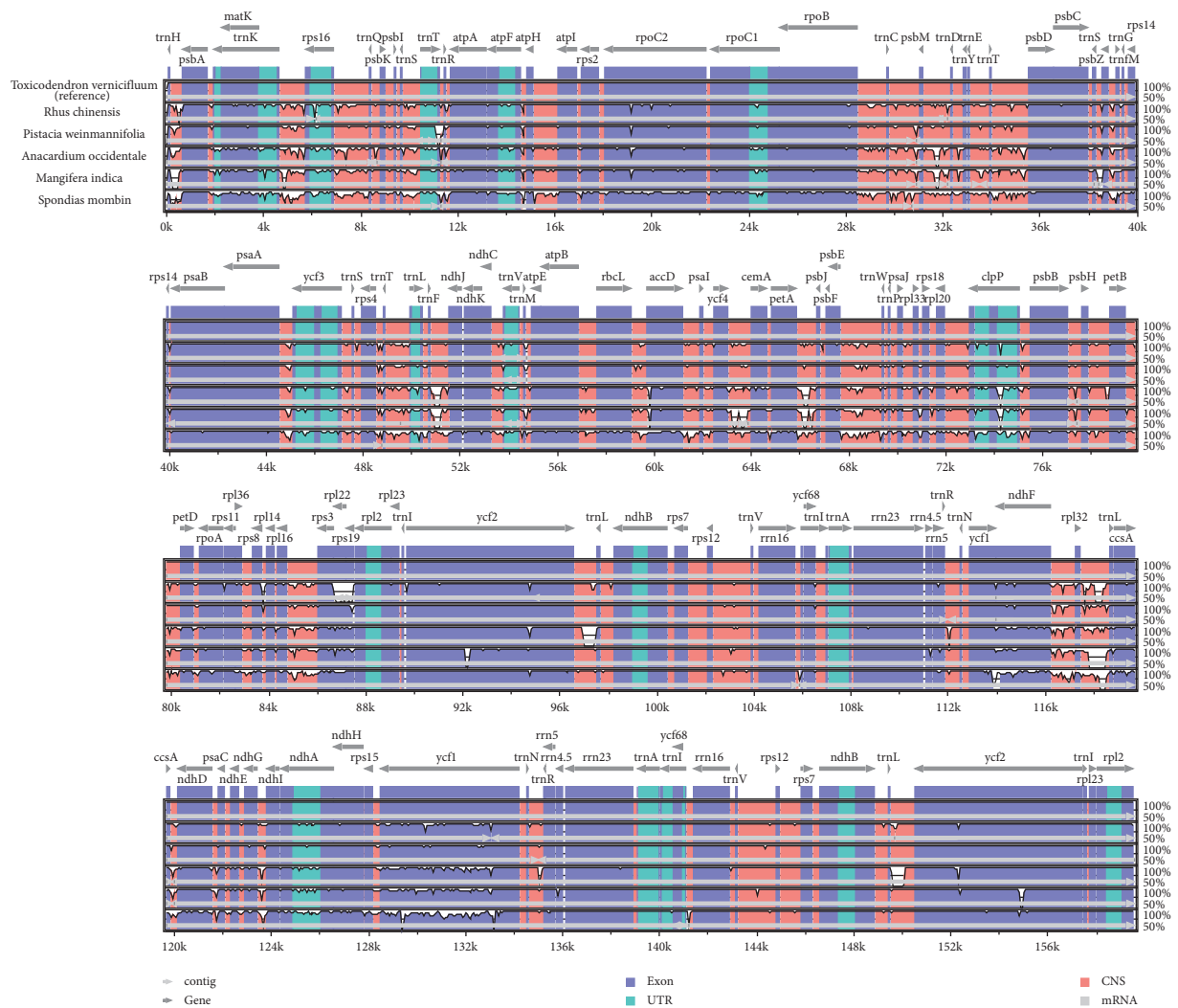


FIGURE 2: Visualization of alignment of the six Anacardiaceae species using mVISTA, with *T. vernicifluum* as the reference. The horizontal axis indicates the coordinates within the chloroplast genome. The vertical scale represents the percentage of identity, ranging from 50% to 100%. The gray arrows above the alignment indicate the genes' orientations.

with normal gene *ndhF* in the IRA/SSC border for many higher plants [67, 68]. This was observed for *T. vernicifluum*, *R. chinensis*, and *S. mombin*, but not for *P. weinmannifolia*, *A. occidentale*, and *M. indica*.

3.6. Phylogenetic Inference. Using HomBlocks, nine locally collinear blocks (LCBs) were identified across the cp genome sequences of 52 species in Sapindales, and two outgroups from Brassicales and Huerteales. After trimming, the final alignment produced a matrix of 65,219 bp (Figure S3), including 20,974 variable sites and 13,581 parsimony-informative sites. Most protein-coding genes have been integrated into the final alignment (Figure S4). The ML and Bayesian analyses yielded identical tree topologies, with all the nodes at or above the generic level allocated posterior probabilities (PP) ≥ 0.99 and bootstrap support (BS) values $\geq 80\%$ (except for the genus *Dipteronia*, Figure 6). Our results showed that Anacardiaceae and

Burseraceae form a monophyletic clade (1.0 PP/100 BS), and the remaining four families form another clade (0.998 PP/83 BS). Simaroubaceae and Rutaceae were resolved as sister to Meliaceae with strong support (1.0 PP/100 BS). Within the Anacardiaceae, the genus *Spondias* was found to be sister to the rest five genera (1.0 PP/100 BS), and *T. vernicifluum* was sister to the clade of *R. chinensis* and two *Pistacia* species (1.0 PP/100 BS).

3.7. Differences of Plastomes between Diploid and Triploid *T. vernicifluum*. The size of the cp genome (159,571 bp) was completely consistent with the result of *T. vernicifluum* cv. *Dahongpao*, a natural triploid cultivar of the species [39]. Only two single nucleotide variants were detected between the two genotypes: one located at the *rpoB-trnC*(GCA) spacer and another within the *ycf1* gene region (Table S8). No indels were observed between them.

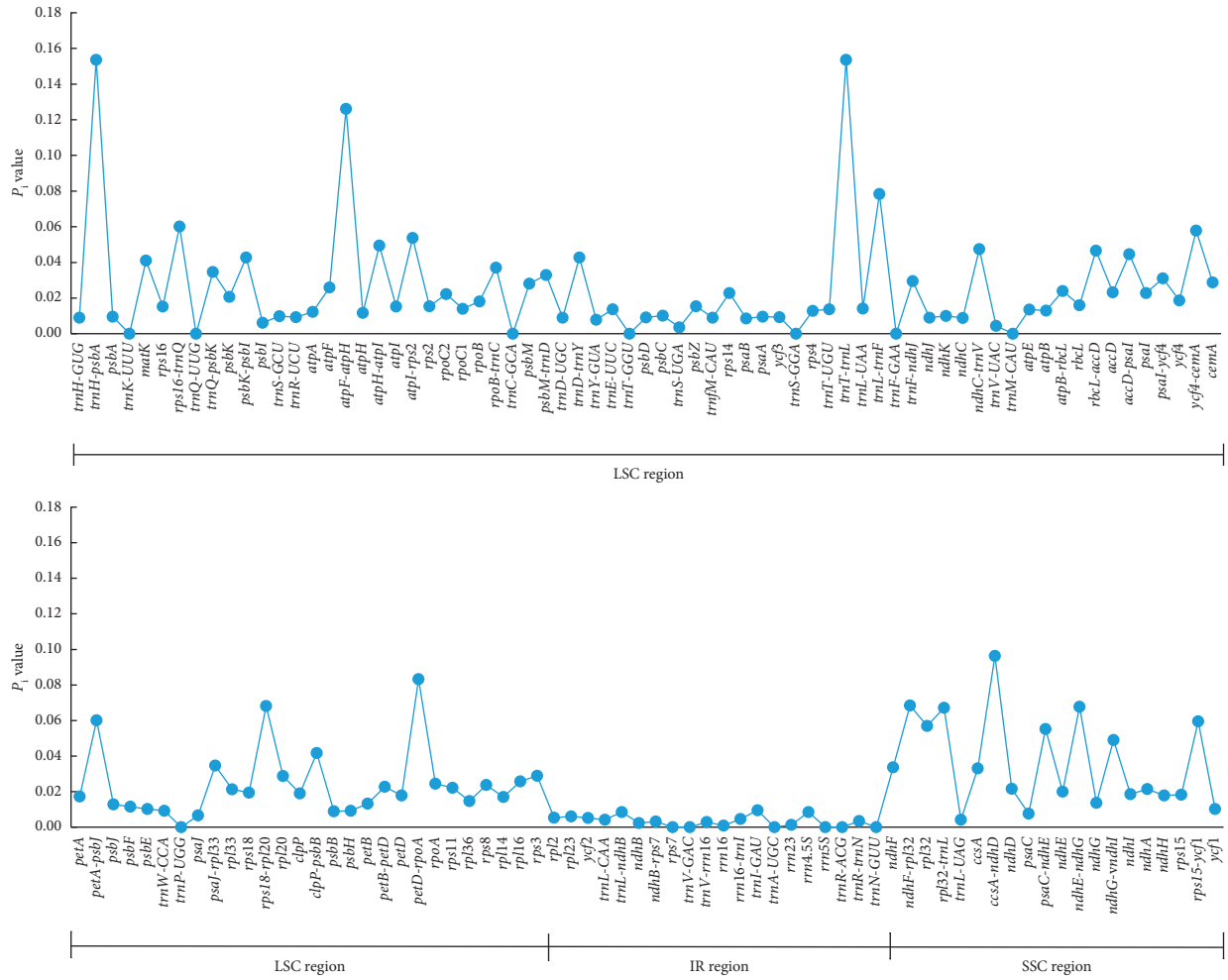


FIGURE 3: Nucleotide diversity (P_i) values among the six Anacardiaceae species.

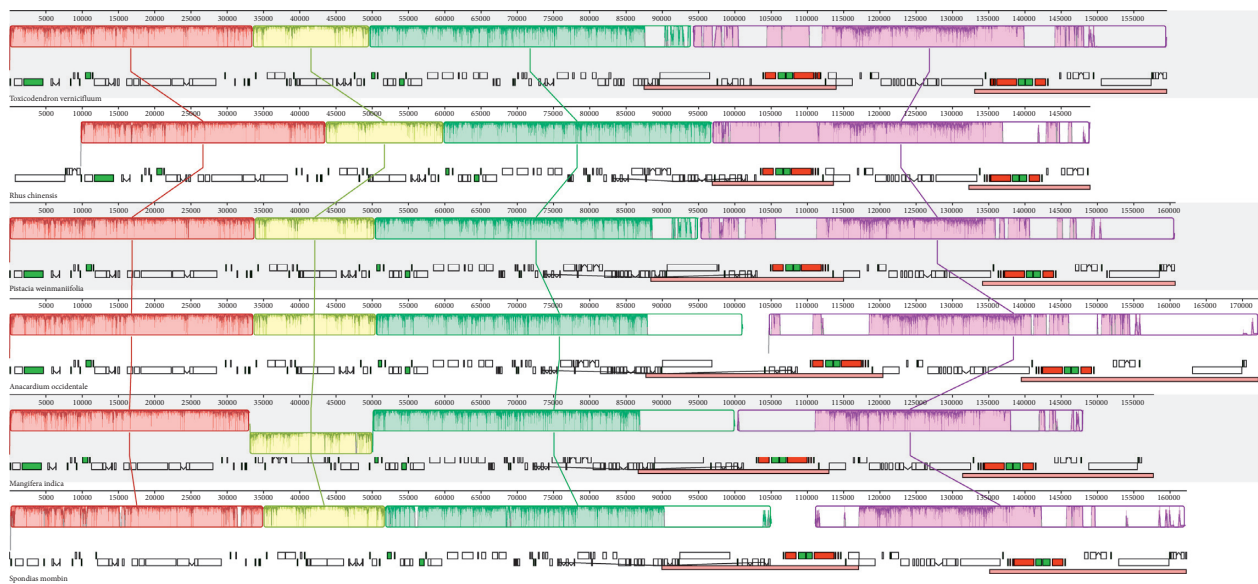


FIGURE 4: Gene map and MAUVE alignment of six Anacardiaceae chloroplast genomes. Within each of the alignment, local collinear blocks are represented by blocks of the same color connected by lines.

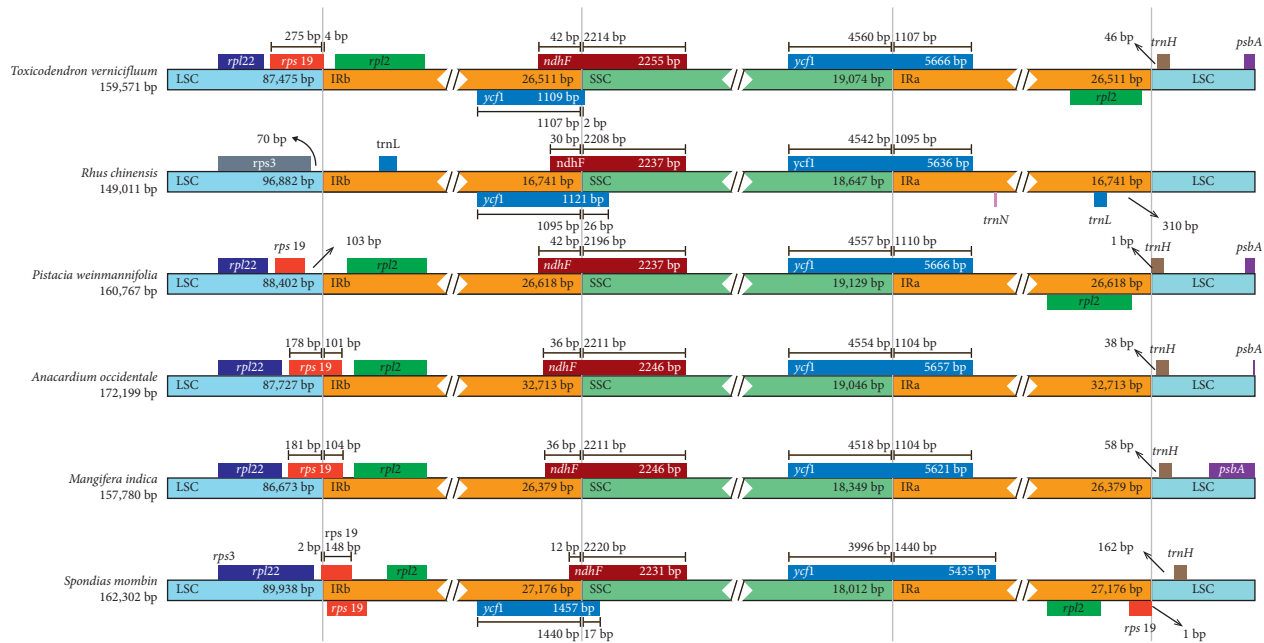


FIGURE 5: Comparison of IR-SC border positions across chloroplast genomes of six Anacardiaceae species. Genes are denoted by colored boxes. The gaps between the genes and the boundaries are indicated by the base lengths (bp).

4. Discussion

In this study, we sequenced the complete cp genome of the Chinese lacquer tree by using Illumina high-throughput sequencing technology. No size variation and only two single nucleotide variants were detected between diploid and triploid of *T. vernicifluum*, suggesting that the species exhibited extremely conserved genome size and structure at the intraspecific level. Compared with closely related species in Anacardiaceae, the length of the cp genome of *T. vernicifluum* is shorter than those of *A. occidentale* (172,199 bp), *S. mombin* (162,302 bp), and *P. weinmannifolia* (160,767 bp), but longer than those of *M. indica* (157,780 bp) and *R. chinensis* (149,011 bp) (Table 1). We found that the total length of the complete cp genome was significantly correlated with the length of IR ($R=0.962$, $P=0.002$) (Figure S2), suggesting that the IR expansion and contraction may be a critical factor underlying the size variation of cp genomes [69–72]. The shortest length of the cp genome of *R. chinensis* was mainly attributed to the loss of a long fragment (~9,800 bp in length) at the IRb/LSC boundary, which contains four genes including *rpl2*, *rpl23*, *trnI-CAU*, and *ycf2*. Previous studies have shown that the loss of a small number of genes, or even all, of the IR is more common in conifer species from cupressophytes and Pinaceae [73, 74]. Nonetheless, similar cases were still reported for several angiosperms, such as *Erodium* L'Hér., *Geranium* L., *Monsonia* L. [75], *Trifolium* L. [76], and *Vicia* L. [77]. For those species, the loss of the complete or partial IR has shifted numerous genes, e.g., *rpl2*, *rpl23*, and *ndhB*, into the SC regions. Recent studies have revealed that most genes that moved from the IR into the SC exhibited higher synonymous substitution rates consistent with the SC genes, and IR localization is a critical factor underlying the reduced

substitution rates in plant plastomes [78]. Among the six Anacardiaceae species, the longest length of the cp genome was observed for *A. occidentale*, which was associated with the size expansion of *ycf2-trnL* spacer (~6,200 bp longer than other five species) in the IR regions. Based on experimental evidence, Rabah et al. [79] have shown that an intracellular gene transfer event of mitochondrial DNA into the plastome may have occurred <20 million years ago in a single clade of the genus *Anacardium*. This event led to an insertion of an ~6,700 bp fragment between *ycf2* and *trnL-CAA*, which is also responsible for the highly expanded IR of *A. occidentale*.

We compared the complete cp genome structure of *T. vernicifluum* with five species from different genera in Anacardiaceae. Our results indicated that the six Anacardiaceae cp genomes were relatively conserved (Figure 4), which is consistent with the slow rates of sequence and structural evolution of plant plastomes [78, 80, 81]. However, the synteny analysis still detected a large inverted segment (16,910 bp in length) in the LSC region of *M. indica*. Similar events have been reported for other species of Malvids, such as *Aquilaria sinensis* (~16,000 bp in length, from *rpl20* to *rbcl* genes) [82]. Previous studies have shown that tRNA activity may be a key factor triggering the inversion events [83]. We found that both the upper and lower flanks of the inverted segment identified in this study had tRNA genes, i.e., *trnE-UUC* and *trnT-UGU* at the upper flank and *trnT-GGU* and *trnL-UAA* at the lower flank. Furthermore, higher genetic variation and lower GC content of the flank regions could also promote gene rearrangements in plastid genomes [84].

The mVISTA analysis showed that intergenic spacers were more variable than genes (Figure 2). Numerous divergence hotspot regions (e.g., *trnH-psbA*, *trnT-trnL*, *atpF-atpH*, *ccsA-ndhD*, *petD-rpoA*, and *trnL-trnF*) were identified

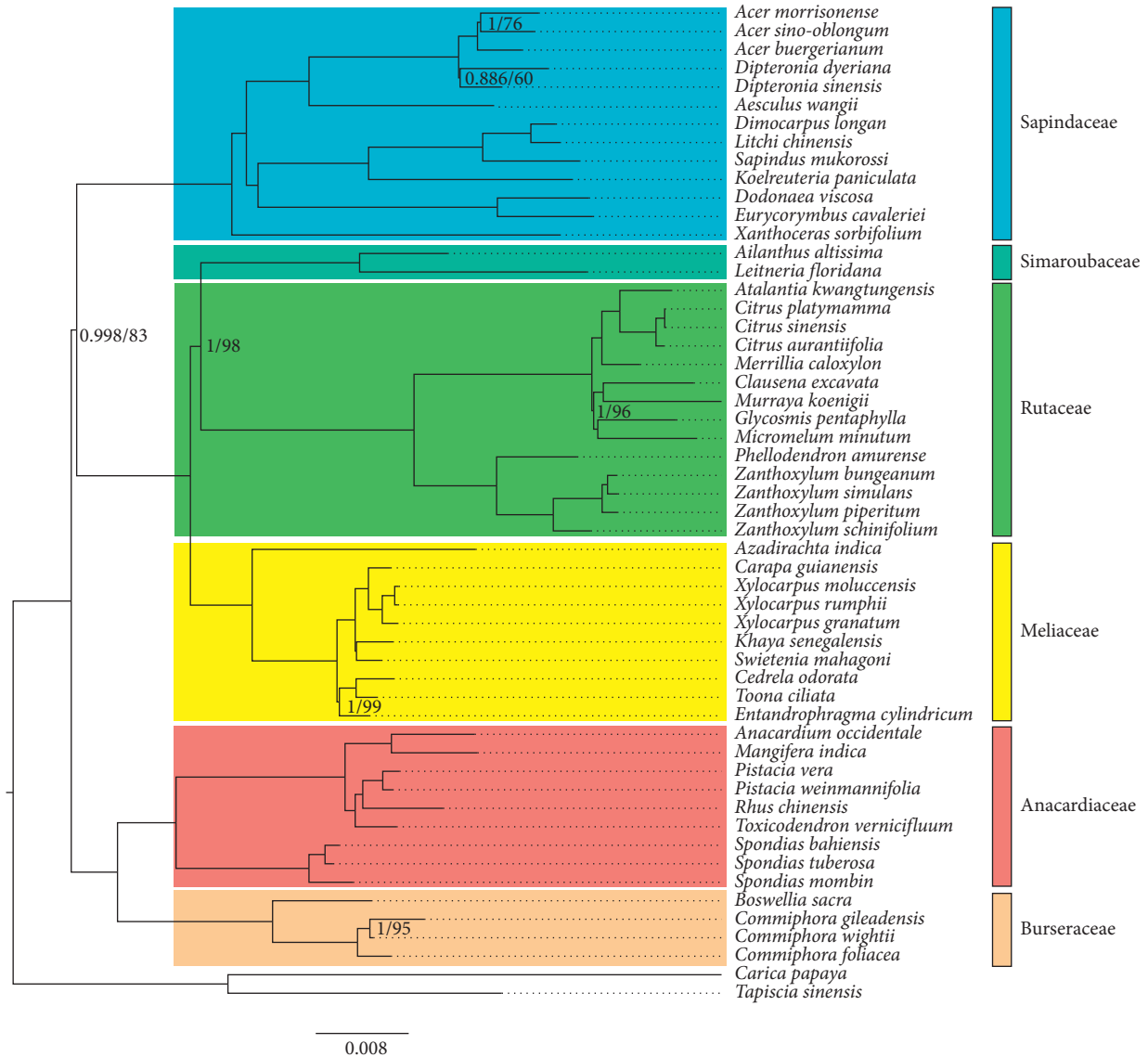


FIGURE 6: Bayesian-inference (BI) and maximum likelihood (ML) analyses based on the chloroplast genome sequences of 52 species from six families in Sapindales and two outgroups of Brassicales and Huerteales. Above each node, the first number indicates the Bayesian posterior probability (PP), and the second number indicates the ML bootstrap value (BS). Nodes with posterior probability of 1 and bootstrap value of 100 are not labeled.

through calculating and comparing the nucleotide diversity (P_i) values (Figure 3). Those highly variable loci can be used as potential molecular markers for the phylogenetic studies in the future [85, 86]. Among those, *trnL-trnF* has been proved to be a useful marker for the phylogenetic studies of *Toxicodendron*. A total of 82 variable sites and 45 parsimony-informative sites were detected across 68 accessions of *Toxicodendron* and some closely related taxa [28]. A more detailed study identified 180 variable sites among the *trnL-trnF* sequences of 85 taxa, representing 22 of the 30 species in *Toxicodendron* and 60 related taxa in Anacardiaceae [87]. This marker was also found to be useful in analyzing the intraspecific variation of *T. vernicifluum*. Combined with the sequence of *trnL* intron, Suzuki et al. [3] have detected three chloroplast DNA haplotypes across populations from China,

South Korea, and Japan. Notably, only one of them was shared by populations in northeastern China, South Korea, and Japan, suggesting that the Chinese lacquer trees in Japan were more likely to be introduced from Liaoning and Shandong provinces of China.

The phylogenetic relationships among several families of Sapindales, such as Sapindaceae, Simaroubaceae, Rutaceae, and Meliaceae, are not fully resolved in the Angiosperm Phylogeny Website (<http://www.mobot.org/MOBOT/research/APweb/>). A recent phylogenetic study based on plastid *rbcl*, *atpB*, and *trnL-trnF* sequences indicated that Simaroubaceae was sister to Meliaceae, with moderate support (0.98 PP/82 BS), but the position of Sapindaceae could not be resolved with confidence [30]. Using 80 genes of plastomes, Lin et al. [31] found that Simaroubaceae was

sister to Rutaceae with strong support (100 BS), but the position of Sapindaceae was still poorly supported (57 BS). In this study, both BI and ML analyses generated a well-resolved phylogeny of the six families of Sapindales. Two distinct clades were recognized: one consisting of Anacardiaceae and Burseraceae (1.0 PP/100 BS) and the other comprised of the remaining four families (0.998 PP/83 BS). In addition, our results strongly supported that Simaroubaceae and Rutaceae form a monophyletic group as sister to Meliaceae (1.0 PP/100 BS). These findings were consistent with the results of the angiosperm phylogeny based on five plastid and nuclear markers [32] and also with the topology of the most recent plastid phylogenomic angiosperm (PPA) tree [88]. Notably, all of those studies using different dataset presented short internodes (e.g., at the base of the Sapindaceae + Simaroubaceae + Rutaceae + Meliaceae clade, and at the base of the Simaroubaceae + Rutaceae + Meliaceae clade) connected by long branches, indicating that rapid radiation may have occurred among those families [89, 90].

Within the Anacardiaceae, both BI and ML analyses supported that *T. vernicifluum* was more closely related to *R. chinensis* and two *Pistacia* species. This is consistent with the phylogeny obtained through three nuclear DNA (ITS, ETS, and NIA-i3) and two chloroplast (*ndhF* and *trnL-trnF*) regions [87]. However, due to lack of cp genome data for congeneric species, we did not provide more details on the intrageneric relationships of *Toxicodendron*. More studies on the plastomes of those species are expected to provide new insight into the evolutionary history of *T. vernicifluum* and its siblings.

5. Conclusions

In this study, we assembled, annotated, and analyzed the cp genome of *T. vernicifluum*, an important commercial arbor species widely cultivated in East Asia for producing highly durable lacquer. Forty-nine mononucleotide microsatellites, one dinucleotide microsatellite, two complex microsatellites, and 49 long repeats were determined. Several hotspots (e.g., *trnH-psbA*, *trnT-trnL*, *atpF-atpH*, *ccsA-ndhD*, *petD-rpoA*, and *trnL-trnF*) of intergeneric divergence were also identified. The unique inversion (in *M. indica*), insertion (in *A. occidentale*), and gene loss (in *R. chinensis*) events may provide informative markers for phylogenetic resolution among different genera in Anacardiaceae. Both BI and ML analyses revealed that the genus *Toxicodendron* is closely related to *Pistacia* and *Rhus*. The phylogenetic relationships of the six families in Sapindales were well resolved, strongly supporting the topology that the clade including Simaroubaceae and Rutaceae is sister to Meliaceae. The genomic resources presented in this study will be useful for further studies on evolutionary patterns of *T. vernicifluum* and its closely related species.

Data Availability

The complete chloroplast genome sequence can be found in GenBank with accession no. MK419151. The data used to

support the findings of this study are included within the supplementary information files.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the National Key Research and Development Program of China (no. 2017YFD060130501) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors thank Genepioneer Biotechnologies Co. Ltd. for sample processing and Dr. Cheng Zhang from Nanjing Forestry University for his assistance with data analyses.

Supplementary Materials

Figure S1: gene map and MAUVE alignment of five Anacardiaceae chloroplast genomes with *Rhus chinensis* removed. Figure S2: the linear correlation between the length of IR and the total length of the complete chloroplast genome sequence. Figure S3: the final alignment produced by the HomBlocks pipeline. Figure S4: visualization of genes that were integrated into the final alignment and their corresponding regions. Table S1: GenBank accession numbers of the complete chloroplast genome sequences of 52 species in Sapindales and two outgroups from Brassicales and Huerteales used for the phylogenetic analyses. Table S2: the best-fit partitioning schemes and DNA substitution models determined by PartitionFinder. Table S3: genes contained in the *Toxicodendron vernicifluum* chloroplast genome. Table S4: genes with introns in the *Toxicodendron vernicifluum* chloroplast genome. Table S5: the codon number and relative synonymous codon usage (RSCU) values calculated based on the coding sequences of 81 protein-coding genes in the complete chloroplast genome of *Toxicodendron vernicifluum*. Table S6: simple sequence repeats (SSRs) of the *Toxicodendron vernicifluum* chloroplast genome. Table S7: long repeats in the *Toxicodendron vernicifluum* chloroplast genome. Table S8: two single nucleotide variants between the complete chloroplast genome of *Toxicodendron vernicifluum* and *T. vernicifluum* cv. *Dahongpao*. (Supplementary Materials)

References

- [1] K. Hashida, M. Tabata, K. Kuroda et al., "Phenolic extractives in the trunk of *Toxicodendron vernicifluum*: chemical characteristics, contents and radial distribution," *Journal of Wood Science*, vol. 60, no. 2, pp. 160–168, 2014.
- [2] D. Gledhill, *The Names of Plants*, Cambridge University Press, Cambridge, UK, 2008.
- [3] M. Suzuki, S. Noshiro, and T. Tanaka, "Origin of urushi (*Toxicodendron vernicifluum*) in the neolithic jomon period of Japan," *Bulletin of the National Museum of Japanese History*, vol. 187, pp. 49–70, 2014.
- [4] M. Suzuki, K. Yonekura, and S. Noshiro, "Distribution and habitat of *Toxicodendron vernicifluum* (Stokes) F. A. Barkley.

- (Anacardiaceae) in China,” *Japan Journal of Historic Botany*, vol. 15, no. 1, pp. 58–62, 2007.
- [5] S. Noshiro and M. Suzuki, “*Rhus verniciflua* Stokes grew in Japan since the early jomon period,” *Japanese Journal of Historic Botany*, vol. 12, no. 1, pp. 3–11, 2004.
- [6] M. Zhao, C. Liu, G. Zheng, S. Wei, and Z. Hu, “Comparative studies of bark structure, lacquer yield and urushiol content of cultivated *Toxicodendron vernicifluum* varieties,” *New Zealand Journal of Botany*, vol. 51, no. 1, p. 13, 2013.
- [7] D. H. Hong, S. B. Han, C. W. Lee et al., “Cytotoxicity of urushiols isolated from sap of Korean lacquer tree (*Rhus vernicifera* Stokes),” *Archives of Pharmacal Research*, vol. 22, no. 6, pp. 638–641, 1999.
- [8] N. Cho, J. H. Choi, and H. Yang, “Neuroprotective and anti-inflammatory effects of flavonoids isolated from *Rhus verniciflua* in neuronal HT22 and microglial BV2 cell lines,” *Food and Chemical Toxicology*, vol. 50, no. 6, pp. 1940–1945, 2012.
- [9] S. Wang, B. Xie, H. Fang et al., “Oil content and fatty acid composition of *Toxicodendron vernicifluum* seeds from the Qinling Mountains,” *Journal of Central South University of Forestry & Technology*, vol. 31, no. 3, pp. 97–101, 2011.
- [10] K. H. Kim, E. Moon, S. U. Choi, C. Pang, S. Y. Kim, and K. R. Lee, “Identification of cytotoxic and anti-inflammatory constituents from the bark of *Toxicodendron vernicifluum* (Stokes) F. A. Barkley,” *Journal of Ethnopharmacology*, vol. 162, pp. 231–237, 2015.
- [11] Y. Xie, J. Zhang, W. Liu, N. Xie, F. Feng, and W. Qu, “New urushiols with platelet aggregation inhibitory activities from resin of *Toxicodendron vernicifluum*,” *Fitoterapia*, vol. 112, pp. 38–44, 2016.
- [12] Z. Meng, S. Wei, and Z. Hu, “Tylosoids in latexiferous canals of *Toxicodendron vernicifluum*,” *Pakistan Journal of Botany*, vol. 46, no. 3, pp. 1039–1045, 2014.
- [13] M. Zhao and C. Hu, “An ultrastructural study of the development of resin canals and lacquer secretion in *Toxicodendron vernicifluum* (Stokes) F. A. Barkley,” *South African Journal of Botany*, vol. 116, pp. 61–66, 2018.
- [14] M. Zhao and H. Zhu, “Development and morphology of stone cells in phloem of *Toxicodendron vernicifluum*,” *Trees*, vol. 28, no. 5, pp. 1553–1558, 2014.
- [15] Y. Hiraoka, S. Hanaoka, A. Watanabe, T. Kawahara, and M. Tabata, “Evaluation of the growth traits of *Toxicodendron vernicifluum* progeny based on their genetic groups assigned using new microsatellite markers,” *Silvae Genetica*, vol. 63, no. 6, pp. 267–274, 2014.
- [16] Y. Li, X. Zhang, Y. Yang, and B. Duan, “Soil cadmium toxicity and nitrogen deposition differently affect growth and physiology in *Toxicodendron vernicifluum* seedlings,” *Acta Physiologiae Plantarum*, vol. 35, no. 2, pp. 529–540, 2013.
- [17] L. Li, K. Cui, S. X. Liao et al., “Seed biological characteristics of *Toxicodendron vernicifluum* in different provenances,” *Southwest China Journal of Agricultural Sciences*, vol. 29, no. 5, pp. 1219–1224, 2016.
- [18] X. P. Zhao and S. N. Wei, “Genetic evaluation of *Toxicodendron vernicifluum* cultivars using amplified fragment length polymorphism markers,” *Journal of Molecular Cell Biology*, vol. 40, no. 4, pp. 262–266, 2007.
- [19] S. Wei, X. Zhao, M. Tian, L. Li, and Z.-h. Hu, “Application of plant morphology and AFLP molecular markers to identify *Toxicodendron vernicifluum* varieties of Shaanxi,” *Acta Botanica Borealis-Occident Sinica*, vol. 30, no. 4, pp. 665–671, 2010.
- [20] P. Zhang, S. Liao, K. Cui et al., “Resources and species of *Toxicodendron* and its industry development prospects in China,” *World Forestry Research*, vol. 26, no. 2, pp. 65–69, 2013.
- [21] J. Zhang, “Review and prospect of research on *Toxicodendron vernicifluum* in China in the past 35 years,” *Journal of Chinese Lacquer*, vol. 3, no. 4, pp. 7–14, 1984.
- [22] Z. Feng, “A preliminary study on the classification of *Toxicodendron vernicifluum* cultivars,” *Journal of Chinese Lacquer*, vol. 3, no. 3, pp. 1–5, 1984.
- [23] D. D. Vu, T. T.-X. Bui, and T. H.-N. Nguyen, “Isolation and characterization of polymorphic microsatellite markers in *Toxicodendron vernicifluum*,” *Czech Journal of Genetics and Plant Breeding*, vol. 54, no. 1, pp. 17–25, 2018.
- [24] T. T. Bui, D. D. Vu, Q. H. Dang et al., “High genetic diversity and population structure of lacquer cultivars (*Toxicodendron vernicifluum*) in Shaanxi province, China revealed by SSR markers,” *Research Journal of Biotechnology*, vol. 12, no. 10, pp. 14–23, 2017.
- [25] S. Walker, J. Williams, J. Lear, and M. Beck, “FS11.5 *Toxicodendron dermatitis* in the United Kingdom,” *Contact Dermatitis*, vol. 50, no. 3, p. 163, 2008.
- [26] H.-L. Li, “Floristic relationships between eastern Asia and eastern North America,” *Transactions of the American Philosophical Society*, vol. 42, no. 2, pp. 371–429, 1952.
- [27] T. L. Min and A. Barfod, *Flora of China: Anacardiaceae*, p. 335, Science Press, Beijing, China, 1980.
- [28] Z.-L. Nie, H. Sun, Y. Meng, and J. Wen, “Phylogenetic analysis of *Toxicodendron* (Anacardiaceae) and its biogeographic implications on the evolution of north temperate and tropical intercontinental disjunctions,” *Journal of Systematics and Evolution*, vol. 47, no. 5, pp. 416–430, 2009.
- [29] M. J. M. Christenhusz and J. W. Byng, “The number of known plant species in the world and its annual increase,” *Phytotaxa*, vol. 261, no. 3, pp. 201–217, 2016.
- [30] A. N. Muellner-Riehl, A. Weeks, J. W. Clayton et al., “Molecular phylogenetics and molecular clock dating of Sapindales based on plastid *rbcL*, *atpB* and *trnL-trnF* DNA sequences,” *Taxon*, vol. 65, no. 5, pp. 1019–1036, 2016.
- [31] N. Lin, M. J. Moore, T. Deng et al., “Complete plastome sequencing from *Toona* (Meliaceae) and phylogenomic analyses within Sapindales,” *Applications in Plant Sciences*, vol. 6, no. 4, Article ID e1040, 2018.
- [32] S. Magallón, S. Gómez-Acevedo, L. L. Sánchez-Reyes, and T. Hernández-Hernández, “A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity,” *New Phytologist*, vol. 207, no. 2, pp. 437–453, 2015.
- [33] C. W. Birky, “Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution,” *Proceedings of the National Academy of Sciences*, vol. 92, no. 25, pp. 11331–11338, 1995.
- [34] W. Dong, C. Xu, T. Cheng, K. Lin, and S. Zhou, “Sequencing angiosperm plastid genomes made easy: a complete set of universal primers and a case study on the phylogeny of Saxifragales,” *Genome Biology and Evolution*, vol. 5, no. 5, pp. 989–997, 2013.
- [35] H. Daniell, C.-S. Lin, M. Yu, and W.-J. Chang, “Chloroplast genomes: diversity, evolution and applications in genetic engineering,” *Genome Biology*, vol. 17, no. 1, p. 134, 2016.
- [36] L. Liu, Y. Wang, and P. He, “Chloroplast genome analyses and genomic resource development for epilithic sister genera *Oresitrophe* and *Mukdenia* (Saxifragaceae), using genome skimming data,” *BMC Genomics*, vol. 19, no. 1, pp. 235–252, 2018.

- [37] C. Alkan, S. Sajjadian, and E. E. Eichler, "Limitations of next-generation genome sequence assembly," *Nature Methods*, vol. 8, no. 1, pp. 61–65, 2011.
- [38] Y. Zhong, D. Zong, A. Zhou, X. He, and C. He, "The complete chloroplast genome of the *Toxicodendron vernicifluum* cv. Dahongpao, an elite natural triploid lacquer tree," *Mitochondrial DNA Part B*, vol. 4, no. 1, pp. 1227–1228, 2019.
- [39] J. J. Doyle and J. L. Doyle, "A rapid DNA isolation procedure for small quantities of fresh leaf tissue," *Phytochemical Bulletin*, vol. 19, pp. 11–15, 1987.
- [40] N. Dierckx, P. Mardulyn, and G. Smits, "NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data," *Nucleic Acids Research*, vol. 45, no. 4, p. e18, 2017.
- [41] C. Liu, L. Shi, Y. Zhu et al., "CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences," *BMC Genomics*, vol. 13, no. 1, p. 715, 2012.
- [42] T. M. Lowe and S. R. Eddy, "tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence," *Nucleic Acids Research*, vol. 25, no. 5, pp. 955–964, 1997.
- [43] M. Lohse, O. Drechsel, and R. Bock, "OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes," *Current Genetics*, vol. 52, no. 5–6, pp. 267–274, 2007.
- [44] D. Zhang, F. Gao, W. X. Li et al., "PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies," *Molecular Ecology Resources*, vol. 20, no. 1, pp. 348–355, 2020.
- [45] T. Thiel, W. Michalek, R. Varshney, and A. Graner, "Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.)," *Theoretical and Applied Genetics*, vol. 106, no. 3, pp. 411–422, 2003.
- [46] S. Kurtz, J. V. Choudhuri, E. Ohlebusch, C. Schleiermacher, J. Stoye, and R. Giegerich, "REPuter: the manifold applications of repeat analysis on a genomic scale," *Nucleic Acids Research*, vol. 29, no. 22, pp. 4633–4642, 2001.
- [47] K. A. Frazer, L. Pachter, A. Poliakov, E. M. Rubin, and I. Dubchak, "Computational tools for comparative genomics," *Nucleic Acids Research*, vol. 32, no. S2, pp. W273–W279, 2004.
- [48] A. E. Darling, B. Mau, and N. T. Perna, "ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement," *PLoS One*, vol. 5, no. 6, Article ID e11147, 2010.
- [49] A. Amiryousefi, J. Hyvönen, and P. Poczai, "IRscope: an online program to visualize the junction sites of chloroplast genomes," *Bioinformatics*, vol. 34, no. 17, pp. 3030–3031, 2018.
- [50] K. Katoh and D. M. Standley, "MAFFT multiple sequence alignment software version 7: improvements in performance and usability," *Molecular Biology and Evolution*, vol. 30, no. 4, pp. 772–780, 2013.
- [51] J. Rozas, A. Ferrer-Mata, J. C. Sánchez-DelBarrio et al., "DnaSP 6: DNA sequence polymorphism analysis of large data sets," *Molecular Biology and Evolution*, vol. 34, no. 12, pp. 3299–3302, 2017.
- [52] B. Bremer, K. Bremer, M. W. Chase et al., "An update of the Angiosperm Phylogeny group classification for the orders and families of flowering plants: APG IV," *Botanical Journal of the Linnean Society*, vol. 161, no. 2, pp. 105–121, 2009.
- [53] G. Bi, Y. X. Mao, Q. K. Xing, and M. Cao, "HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching," *Genomics*, vol. 110, no. 1, pp. 18–22, 2017.
- [54] A. C. E. Darling, B. Mau, F. R. Blattner, and N. T. Perna, "Mauve: multiple alignment of conserved genomic sequence with rearrangements," *Genome Research*, vol. 14, no. 7, pp. 1394–1403, 2004.
- [55] J. Castresana, "Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis," *Molecular Biology and Evolution*, vol. 17, no. 4, pp. 540–552, 2002.
- [56] R. Lanfear, B. Calcott, S. Y. W. Ho, and S. Guindon, "PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses," *Molecular Biology and Evolution*, vol. 29, no. 6, pp. 1695–1701, 2012.
- [57] F. Ronquist, M. Teslenko, P. van der Mark et al., "MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space," *Systematic Biology*, vol. 61, no. 3, pp. 539–542, 2012.
- [58] D. Silvestro and I. Michalak, "raxmlGUI: a graphical front-end for RAxML," *Organisms Diversity & Evolution*, vol. 12, no. 4, pp. 335–337, 2012.
- [59] A. Stamatakis, "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies," *Bioinformatics*, vol. 30, no. 9, pp. 1312–1313, 2014.
- [60] G.-L. Xin, J.-q. Liu, and J. Liu, "Complete chloroplast genome of an endangered tree species, *Toona ciliata* (Sapindales: Meliaceae)," *Mitochondrial DNA Part B*, vol. 3, no. 2, pp. 663–664, 2018.
- [61] A. Khan, S. Asaf, A. L. Khan et al., "First complete chloroplast genomics and comparative phylogenetic analysis of *Commiphora gileadensis* and *C. foliacea*: myrrh producing trees," *PLoS One*, vol. 14, no. 1, Article ID e0208511, 2019.
- [62] S.-Y. Chen and X.-Z. Zhang, "Characterization of the complete chloroplast genome of *Xanthoceras sorbifolium*, an endangered oil tree," *Conservation Genetics Resources*, vol. 9, no. 4, pp. 595–598, 2017.
- [63] H. Dashnow, S. Tan, D. Das, S. Eastal, and A. Oshlack, "Genotyping microsatellites in next-generation sequencing data," *BMC Bioinformatics*, vol. 16, no. S2, p. A5, 2015.
- [64] S. H. Park, I. P. Sang, J. Gil et al., "Development of chloroplast SSR markers to distinguish *Codonopsis* species," *Korean Journal of Horticultural Science & Technology*, vol. 5, pp. 207–208, 2017.
- [65] M. Chmielewski, K. Meyza, I. Chybicki, A. Dzialuk, M. Litkowiec, and J. Burczyk, "Chloroplast microsatellites as a tool for phylogeographic studies: the case of white oaks in Poland," *iForest—Biogeosciences and Forestry*, vol. 8, no. 6, pp. 765–771, 2015.
- [66] J. Qian, J. Song, H. Gao et al., "The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*," *PLoS One*, vol. 8, Article ID e5760, 2013.
- [67] J. Ma, B. Yang, W. Zhu, L. Sun, J. Tian, and X. Wang, "The complete chloroplast genome sequence of *Mahonia bealei* (Berberidaceae) reveals a significant expansion of the inverted repeat and phylogenetic relationship with other angiosperms," *Gene*, vol. 528, no. 2, pp. 120–131, 2013.
- [68] K. S. Choi and S. Park, "The complete chloroplast genome sequence of *Aster spathulifolius* (Asteraceae); genomic features and relationship with Asteraceae," *Gene*, vol. 572, no. 2, pp. 214–221, 2015.
- [69] R. Kolodner, K. K. Tewari, and R. C. Warner, "Physical studies on the size and structure of the covalently closed circular chloroplast DNA from higher plants," *Biochimica et Biophysica Acta (BBA)—Nucleic Acids and Protein Synthesis*, vol. 447, no. 2, pp. 144–155, 1976.

- [70] X. Li, H. Gao, Y. Wang et al., "Complete chloroplast genome sequence of *Magnolia grandiflora* and comparative analysis with related species," *Science China Life Sciences*, vol. 56, no. 2, pp. 189–198, 2013.
- [71] K.-J. Kim and H. L. Lee, "Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants," *DNA Research*, vol. 11, no. 4, pp. 247–261, 2004.
- [72] L. A. Raubeson, R. Peery, T. W. Chumley et al., "Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*," *BMC Genomics*, vol. 8, no. 1, p. 174, 2007.
- [73] W. Guo, F. Grewe, A. Cobo-Clark et al., "Predominant and substoichiometric isomers of the plastid genome coexist within *Juniperus* plants and have shifted multiple times during Cupressophyte evolution," *Genome Biology and Evolution*, vol. 6, no. 3, pp. 580–590, 2014.
- [74] C.-S. Wu, C.-P. Lin, C.-Y. Hsu, R.-J. Wang, and S.-M. Chaw, "Comparative chloroplast genomes of Pinaceae: insights into the mechanism of diversified genomic organizations," *Genome Biology and Evolution*, vol. 3, pp. 309–319, 2011.
- [75] M. M. Guisinger, J. V. Kuehl, J. L. Boore, and R. K. Jansen, "Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage," *Molecular Biology and Evolution*, vol. 28, no. 1, pp. 583–600, 2011.
- [76] J. D. Palmer, B. Osorio, J. Aldrich, and W. F. Thompson, "Chloroplast DNA evolution among legumes: loss of a large inverted repeat occurred prior to other sequence rearrangements," *Current Genetics*, vol. 11, no. 4, pp. 275–286, 1987.
- [77] C. Li, Y. Zhao, H. Huang, Y. Ding, Y. Hu, and Z. Xu, "The complete chloroplast genome of an inverted-repeat-lacking species, *Vicia sepium*, and its phylogeny," *Mitochondrial DNA Part B*, vol. 3, no. 1, pp. 137–138, 2018.
- [78] A. Zhu, W. Guo, S. Gupta, W. Fan, and J. P. Mower, "Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates," *New Phytologist*, vol. 209, no. 4, pp. 1747–1756, 2016.
- [79] S. O. Rabah, C. Lee, N. H. Hajrah et al., "Plastome sequencing of ten nonmodel crop species uncovers a large insertion of mitochondrial DNA in cashew," *The Plant Genome*, vol. 10, no. 3, p. 20, 2017.
- [80] Y. Huang, J. Wang, Y. Yang, C. Fan, and J. Chen, "Phylogenomic analysis and dynamic evolution of chloroplast genomes in Salicaceae," *Frontiers in Plant Science*, vol. 8, p. 1050, 2017.
- [81] W. Liu, H. Kong, J. Zhou, P. Fritsch, G. Hao, and W. Gong, "Complete chloroplast genome of *Cercis chuniana* (Fabaceae) with structural and genetic comparison to six species in Caesalpinioideae," *International Journal of Molecular Sciences*, vol. 19, no. 5, p. 1286, 2018.
- [82] J. K. Saina, Z. Z. Li, A. W. Gichira, and Y.-Y. Liao, "The complete chloroplast genome sequence of tree of heaven (*Ailanthus altissima* Mill.) (Sapindales: Simaroubaceae), an important pantropical tree," *International Journal of Molecular Sciences*, vol. 19, no. 4, p. e929, 2018.
- [83] J. Hiratsuka, H. Shimada, R. Whittier et al., "The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals," *Molecular and General Genetics MGG*, vol. 217, no. 2-3, pp. 185–194, 1989.
- [84] H. Liu, J. He, C. Ding et al., "Comparative analysis of complete chloroplast genomes of *Anemoclema*, *Anemone*, *Pulsatilla*, and *Hepatica* revealing structural variations among genera in tribe Anemoneae (Ranunculaceae)," *Frontiers in Plant Science*, vol. 9, p. 1097, 2018.
- [85] W. Wang, S. Chen, and X. Zhang, "Whole-genome comparison reveals heterogeneous divergence and mutation hotspots in chloroplast genome of *Eucommia ulmoides* Oliver," *International Journal of Molecular Sciences*, vol. 19, no. 4, p. 1037, 2018.
- [86] I. Park, S. Yang, W. Kim et al., "Sequencing and comparative analysis of the chloroplast genome of *Angelica polymorpha* and the development of a novel indel marker for species identification," *Molecules*, vol. 24, no. 6, p. 1038, 2019.
- [87] Y. Jiang, M. Gao, Y. Meng, J. Wen, X.-J. Ge, and Z.-L. Nie, "The importance of the North Atlantic land bridges and eastern Asia in the post-Boreotropical biogeography of the Northern Hemisphere as revealed from the poison ivy genus (*Toxicodendron*, Anacardiaceae)," *Molecular Phylogenetics and Evolution*, vol. 139, Article ID 106561, 2019.
- [88] H.-T. Li, T.-S. Yi, L.-M. Gao et al., "Origin of angiosperms and the puzzle of the Jurassic gap," *Nature Plants*, vol. 5, no. 5, pp. 461–470, 2019.
- [89] P.-F. Ma, Y.-X. Zhang, C.-X. Zeng, Z.-H. Guo, and D.-Z. Li, "Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (Poaceae)," *Systematic Biology*, vol. 63, no. 6, pp. 933–950, 2014.
- [90] H. Liu, Z. Su, S. Yu et al., "Genome comparison reveals mutation hotspots in the chloroplast genome and phylogenetic relationships of *Ormosia* species," *BioMed Research International*, vol. 2019, Article ID 7265030, 11 pages, 2019.