Chloroplast genome sequence of triploid *Toxicodendron vernicifluum* and comparative analyses with other lacquer chloroplast genomes

Dan Zong  
Key Laboratory for Forestry Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming

Zhensheng Qiao  
Key Laboratory for Forestry Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming

Jintao Zhou  
Key Laboratory for Forestry Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming

Peiling Li  
Key Laboratory for Forest Genetics and Tree Improvement & Propagation in Universities of Yunnan Province, Southwest Forestry University, Kunming

Peihua Gan  
Key Laboratory for Forestry Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming

Meirong Ren  
Key Laboratory of Biodiversity Conservation in Southwest China, State Forestry Administration, Southwest Forestry University, Kunming

Chengzhong He  
hecz@swfu.edu.cn  
Key Laboratory for Forestry Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming

**Research Article**

**Keywords:** triploid *T. vernicifluum*, chloroplast genome, phylogeny analyses, SSR markers

**Posted Date:** November 17th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-2082751/v1

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

Background: Toxicodendron verniciuum, belonging to the family Anacardiaceae, is an important commercial arbor species, which can provide us with the raw lacquer, an excellent adhesive and painting material used to make lacquer ware. Compared with diploid, triploid lacquer tree has a higher yield of raw lacquer and stronger resistance to stress. Triploid T. verniciuum was a newly discovered natural triploid lacquer tree. However, the taxonomy of triploid T. verniciuum has remained uncertain. Here, we sequenced and analyzed the complete chloroplast (cp) genome of triploid T. verniciuum and compared with T. verniciuum cv. Dahongpao, T. verniciuum cv. Hongpigaobachi, T. verniciuum and T. succedaneum based on chloroplast genome and SSR markers.

Results:

The plastome of triploid T. verniciuum is 158,221 bp in length, including a pair of inverted repeats (IRs) of 26,462 bp, separated by a large single-copy region of 86,951 bp and a small single-copy region of 18,346 bp. In total, 116 unique genes including 82 protein-coding genes, 30 tRNA genes and 4 rRNA genes were identified in the triploid T. verniciuum. After nucleotide substitutions, small inversions were analyzed in the chloroplast genomes, five hotspot regions were found, which could be useful molecular genetic markers for future population genetics. Phylogenetic analyses showed that triploid T. verniciuum was a sister to T. verniciuum cv. Dahongpao and T. verniciuum cv. Hongpigaobachi. Moreover, phylogenetic clustering based on the SSR markers showed that all the individuals of triploid T. verniciuum, T. verniciuum cv. Dahongpao and T. verniciuum cv. Hongpigaobachi in one group, while the individuals of T. verniciuum and T. succedaneum in another group.

Conclusions: The current genomic datasets provide pivotal genetic resources to determine the phylogenetic relationships, variety identification, breeding and resource exploitation, and future genetic diversity-related studies of T. verniciuum.

Background

Polyploidy, that is having multiple sets of chromosomes as a consequence of whole-genome duplication, is common in nature and provides a major mechanism for adaptation and speciation [1–2]. The polyploidy genotypes may lead to differences in morphology, physiology and molecular characteristics, etc [3]. Toxicodendron verniciuum, belonging to the family Anacardiaceae, is a deciduous tree species with a toxic sap [4]. The resin and sap which was extracted from lacquer trees were used as paint for culture assets and lacquer wares, making it a natural material of cultural and social significance [5–7]. In addition, lacquer tree is also used as a food additive, natural dye, or in herbal medicine to improve blood circulation and to prevent blood stasis, while the metabolic extract of the leaves has neuroprotective and anti-inflammatory activity [8]. The economic value of different varieties of lacquer can be judged by lacquer yield, the growth rate of lacquer and genetic diversity [9].

Compared with the diploid lacquer tree, the triploid lacquer tree has a higher yield of raw lacquer and stronger resistance to stress [10]. T. verniciuum cv. Dahongpao (2n = 3x = 45) was the first nature triploid lacquer tree which was discovered at Bashan Mountain in Shaanxi Province, China and now is widely introduced into the planting area of lacquer tree [11–12]. Han et al [13] determined the pseudo-polyploidy of T. verniciuum was a natural triploid lacquer by observation of the stomatal characteristic of leaf lower epidermis, analysis of lower epidermis, flow cytometry and measurement of genome size, and it was named as triploid T. verniciuum. However, as an economic tree species, T. verniciuum has been widely introduced and cultivated [14], which lead to the taxonomy classification was difficult to resolve because of considerable phenotypic variability with overlapping morphologies.

Chloroplast (cp) genomes have assembled notable contributions in diverse plant families, setting evolutionary within phylogenetic clades [15–17], because the lack of recombination and maternal transmission means the chloroplast genome is also useful for tracking source populations [18–20] and chloroplast genomes variation could provide valuable genetic markers for the analysis of polyploids [21]. In addition, the cp genome play an important role in the reconstruction of the green plant phylogeny and understanding the origins of economically important cultivated species and changes that have taken place during domestication [22]. Microsatellites are powerful markers to detect genetic variation, because of their high mutation rate and high polymorphism, allowing us to distinguish between closely related individuals, and thus to assess parentage and kinship relationships [23].

In this study, we reported the characteristics of the complete chloroplast genome sequences of triploid T. verniciuum for the first time and compared with other four Toxicodendron cp genomes to investigate the relationship among chloroplast genome, and SSR was used as a tool to facilitate the assessment of molecular diversity and identify related species. To understand the relationships of the triploid T. verniciuum, we constructed the phylogenetic tree using their fully sequenced chloroplast genome sequenced and SSR markers. The results will provide a theoretical basis for variety identification, breeding and resource exploitation.

Results
Morphological traits of five Toxicodendron accessions

The leaf length, leaflet number, leaflet length and leaflet width were compared among the five accessions (Table 1), the results showed that the leaf length of triploid T. verniciiflum was significantly higher than other four accessions. The leaflet number, leaflet length and leaflet width were significantly lower than T. verniciiflum cv. Dahongpao and T. verniciiflum cv. Hongpigaobachi, and higher than T. verniciiflum and T. succedaneum. The leaf shape index of TZT, DHP and GBC (2.03, 2.11 and 2.18, respectively) were lower than TZG (3.21) and TRB (3.31). In addition, all the morphological traits both showed that there were no significant difference between T. verniciiflum cv. Dahongpao and T. verniciiflum cv. Hongpigaobachi.

Table 1
Leaf morphological characteristics of five Toxicodendron accessions

<table>
<thead>
<tr>
<th>Accessions</th>
<th>Leaf length (±)</th>
<th>Leaflet number (±)</th>
<th>Leaflet length (±)</th>
<th>Leaflet width (±)</th>
<th>Leaf shape index (±)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TZT</td>
<td>33.67 ± 4.19A</td>
<td>8.56 ± 0.92C</td>
<td>11.29 ± 1.52B</td>
<td>5.60 ± 0.62B</td>
<td>2.03 ± 0.17B</td>
</tr>
<tr>
<td>DHP</td>
<td>29.41 ± 5.97B</td>
<td>10.97 ± 1.73AB</td>
<td>12.97 ± 1.83A</td>
<td>6.21 ± 1.00A</td>
<td>2.11 ± 0.20B</td>
</tr>
<tr>
<td>GBC</td>
<td>31.02 ± 6.93AB</td>
<td>10.04 ± 1.78B</td>
<td>14.38 ± 2.31A</td>
<td>6.62 ± 0.54A</td>
<td>2.18 ± 0.35B</td>
</tr>
<tr>
<td>TZG</td>
<td>25.17 ± 4.17C</td>
<td>10.93 ± 2.52AB</td>
<td>9.77 ± 1.81D</td>
<td>3.05 ± 0.42C</td>
<td>3.21 ± 0.37A</td>
</tr>
<tr>
<td>TRB</td>
<td>21.08 ± 3.70D</td>
<td>12.22 ± 2.76A</td>
<td>9.28 ± 1.42D</td>
<td>2.86 ± 0.56C</td>
<td>3.31 ± 0.48A</td>
</tr>
</tbody>
</table>


Features of Toxicodendron accessions chloroplast genome

In this study, all the cpDNAs showed a typical circular tetramerous structure, consisting of a pair of inverted repeats (IRs), a large single copy region (LSC), and a small single copy region (SSC) (Fig. 1). The size of cpDNA and its regions were all similar across different Toxicodendron accessions (Table 2). The cpDNA length of Toxicodendron ranges from 158,221 bp (triploid T. verniciiflum) to 159,710 bp (T. verniciiflum). The size of the IR region ranges from 26,462 bp (triploid T. verniciiflum) to 26,534 bp (T. succedaneum), while the SSC and LSC size varies from 18,346 bp (triploid T. verniciiflum) to 19,074 bp (T. verniciiflum cv. Dahongpao and T. verniciiflum cv. Hongpigaobachi) and from 86,951 bp (triploid T. verniciiflum) to 87,636 bp (T. verniciiflum) (Table 2).

The base composition of cp genome sequence was analyzed and found to be 30.7% A, 19.3% C, 18.6% G and 31.4% Tin triploid T. verniciiflum (Table 3). The overall GC content was 38.0%, which was very close to those of other Toxicodendron, e.g., T. verniciiflum cv. Hongpigaobachi (38.0%), T. verniciiflum (37.9%), T. succedaneum (37.9%), T. verniciiflum cv. Dahongpao (38.0%) [24] and T. verniciiflum (37.9%) [25]. Furthermore, the GC contents are unevenly distributed across regions of the cp genome, which were found 36.1%, 42.9% and 32.7% for the LSC, IR and SSC regions, respectively in triploid T. verniciiflum (Table 2). The triploid T. verniciiflum cpDNA consists of 116 unique genes (Fig. 1 and Table S1), including 82 protein-coding genes, 30 tRNA genes and 4 rRNA genes. In the chloroplast genomes of triploid T. verniciiflum, 14 genes (atpF, ndhA, ndhB, petB, petD, rpl2, rpl16, rpoC1, rps16, trnA-UGC, trnG-UCC, trnL-GAU, trnL-UAA and trnV-GAC) contain one intron and while three genes (clpP, rps12 and ycf3) contain two introns.
Ir Expansion And Contraction

Although cp genomes are highly conserved in terms of genomic structure and size, the IR/SC junction position change caused by expansion and contraction of the IR/SC boundary regions was usually considered as a primary mechanism in creating the length variation of the higher plant cp genomes [26–27]. We investigated the position of genes at the junction regions of five chloroplast genomes: **T. vernicifloum cv.**
Dahongpao, *T. verniciiflum* cv. Hongpigaobachi, triploid *T. verniciiflum, T. verniciiflum* and *T. succedaneum* (Fig. 2). At the LSC/IR junction of five accessions, the *rpl2* gene was duplicated at the IR/SSC junction completely and included in the IR region and the *rpl2* gene in the *T. verniciiflum* cv. Dahongpao and *T. verniciiflum* cv. Hongpigaobachi was shifted by 67 bp from IR to LSC at the LSC/IR border and 66 bp, 88 bp and 103 bp from IR to LSC in the triploid *T. verniciiflum, T. verniciiflum* and *T. succedaneum*, respectively. The *ycf1* gene is located at the IRa/SSC border in the five cp genomes, and the junctions of IRa/SSC located in *ycf1* within the SSC and IRa regions almost had the same length (4560 bp and 1107 bp) except *T. succedaneum* (4560 bp and 1101 bp). The gene *ycf1* in the IRb region and gene *ndhF* in the SSC region interlaced at the IRb/SSC border and *ycf1* in the SSC region was astride the border of SSC/IRa. Gene *ndhF* and *ycf1* in the SSC region extended the same number of bases among the five accessions (42 bp and 1107 bp), except *T. succedaneum* (36 bp and 1101 bp). Therefore, we may infer that the expansion and contraction of IR region in triploid *T. verniciiflum* chloroplast genome is fairly stable compared to that of the other chloroplast genomes in *Toxicodendron*.

### Comparative Analysis Of Genome Structure

To investigate the intergeneric divergence of cp genome sequences, the percentage of identity was plotted for five *Toxicodendron* accessions using mVISTA program with DHP as a reference. The alignment revealed high sequence similarity across the five cp genomes and no rearrangement occurred (Fig. 3), which suggests that they are highly conserved. Non-coding and SC regions exhibit higher divergence levels than coding and IR regions, respectively.

The cp genome sequences of the five *Toxicodendron* accessions were aligned by MAUVE, and DHP was used as a reference to compare the gene orders among these cp genomes (Fig. 4). The results showed that all sequences show perfect synteny conservation with no inversion or rearrangements.

The nucleotide variability (Pi) values of the eight cp genomes were calculated with the DnaSP software. A total of 1252 polymorphic sites were detected and the Pi values ranged from 0.00001 to 0.0060. Five Pi value peaks (*tmK-rps16, ycf4-cemA, psbL-petL, ndhF-ccsA* and *ccsA-ndhA*) were recognized as divergence hotspots (Fig. 5). All hotspots were identified as intergenic spacers, which echoed the finding that the non-coding regions exhibited more variations than the coding regions. Overall, the LSC showed more divergence than the SSC and the IRs. The sequences of these highly variable regions could be developed as barcodes for species identification, phylogenetic analysis, and population genetics research.

We investigated SNPs, the most abundant type of mutation, in the five cp genomes, with *T. verniciiflum* cv. Dahongpao (DHP) as the reference. In the gene-coding regions, we detected two SNPs in the comparative combination of GBC-DHP, including one transition (Ts) and one transversion (Tv) SNPs, as well as 121 (67 Ts and 54 Tv), 227 (114 Ts and 113 Tv) and 102 (46 Ts and 56 Tv) SNPs were detected in the combinations of TZT-DHP, TZG-DHP and TRB-DHP (Table 4). Furthermore, 5 (4 Ts and 1 Tv), 260 (96 Ts and 164 Tv), 462 (181 Ts and 281 Tv) and 126 (10 Ts and 116 Tv) SNPs were detected in noncoding regions among the four comparative combinations, respectively (Table S2).

#### Table 4

<table>
<thead>
<tr>
<th>Treat</th>
<th>Ts</th>
<th>Tv</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A-G</td>
<td>C-T</td>
<td>A-T</td>
</tr>
<tr>
<td>GBC-DHP</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TZT-DHP</td>
<td>31</td>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>TZG-DHP</td>
<td>54</td>
<td>60</td>
<td>7</td>
</tr>
<tr>
<td>TRB-DHP</td>
<td>22</td>
<td>24</td>
<td>6</td>
</tr>
</tbody>
</table>

It has been reported that each small inversion is commonly associated with a hairpin secondary structure in the chloroplast genomes [26, 28]. Small inversions are generally detected by performing pairwise comparisons between sequences of closely related taxa [26]. Seven small inversions were identified in the *Toxicodendron* cp genomes and their inverted repeating flanking sequences formed stem-loop structures (Fig. 6). All the inversions were located in noncoding regions including 6 in space (*ccsA-ndhD, trnS-psbZ, atpF-atpH, tmW-trnP, tmG-trmR* and *tmQ-psbK*) and one in intron regions (*rpl16 intron*). Four inversions (*ccsA-ndhD, trnS-psbZ, tmG-trmR* and *tmQ-psbK*) were specific to one species. For example, the inversion in *ccsA-ndhD and trnS-psbZ* were specific to TZT, whereas inversions in *tmG-trmR* and *tmQ-psbK* occurred in *T. verniciiflum*. This suggests that these inversions are polymorphic in one species.
Phylogenetic Analysis Based On The Chloroplast Complete Genome And Ssr Molecular Markers

To identify the phylogenetic position of triploid *T. verniciiflum* in *Toxicodendron*, we used the 10 cp genomes to phylogenetic analysis. Maximum likelihood (ML) and Bayesian inference (BI) were used to construct phylogenetic tree with *Pistacia weinmannifolia* and *Mangifera indica* as out-groups (Fig. 7). The topologies of the ML and BI trees were nearly identical, which both showed that *Toxicodendron* species formed a monophyletic clade (BS = 100, PP = 1). The eight *Toxicodendron* species were divided into two main clades. Clade I contained four accessions (*T. verniciiflum* cv. Dahongpao, *T. verniciiflum* cv. Hongpigaobachi, *T. verniciiflum* cv. Yanggangdamu, and triploid *T. verniciiflum*) and the results showed that *T. verniciiflum* cv. Dahongpao is closely related to *T. verniciiflum* cv. Hongpigaobachi and sister to triploid *T. verniciiflum*. Clade II consist of *T. succedaneum* 1#, *T. verniciiflum*, *T. sylvestre* and *T. succedaneum* and *T. verniciiflum* can be distinguished with *T. succedaneum*. The phylogenetic tree was very helpful for us to understand the phylogenetic relationship among more *Toxicodendron* species.

Genetic distance among the 15 accessions was calculated according to the software GenoDive (Table S3) [29]. Based on Nei's genetic distance coefficient, a dendrogram was obtained using UPGMA cluster analysis. With this result, two groups could be distinguished. The first group was further divided into two subgroups, three accessions (*T. succedaneum* 1#, *T. succedaneum* 2# and *T. succedaneum* 3#) were included in subgroup 1, and the three accessions of *T. verniciiflum* (*T. verniciiflum* 1#, 2# and 3#) were included in the subgroup II. Group II included three accessions of *T. verniciiflum* cv. Dahongpao, *T. verniciiflum* cv. Hongpigaobachi and triploid *T. verniciiflum*, and *T. verniciiflum* cv. Dahongpao was sister to triploid *T. verniciiflum*, while the three accessions can be distinguished from triploid *T. verniciiflum*, and *T. verniciiflum* cv. Dahongpao (Fig. 8).

Discussion

There are obvious differences in leaf morphology between diploid and polyploidy, which is one of the simpler methods to identify plant ploidy by morphological differences [30]. Leaf shape index (length/width) is an important parameter of leaf shape. The larger the leaf shape index of plant is, the longer and narrower the leaves are, while the smaller the leaf shape index is, the rounder the leaves are [31]. Liu et al. [32] used the observational method and paraffin section technique to analyze the intraspecific differences in four kinds of *Toxicodendron* species and the results showed that the sharp differences presented in height, shape, bark and leaves of the four varieties. In this study, we compared the five foliar traits of five different *Toxicodendron* species; the results showed that the leaf length of triploid *T. verniciiflum*, *T. verniciiflum* cv. Dahongpao and *T. verniciiflum* cv. Hongpigaobachi were significantly higher than *T. verniciiflum* and *T. succedaneum*. In addition, all four morphological characteristics were shown that *T. verniciiflum* cv. Dahongpao and *T. verniciiflum* cv. Hongpigaobachi had no significant difference.

The chloroplast genomes of plants are a valuable resource for developing molecular markers to study intra-species and interspecies evolution [33–34]. The current study showed the first complete cp genome sequence for triploid *T. verniciiflum*, genus *Toxicodendron* and family Anacardiaceae. Further, the cp genome was compared with four cp genomes of related species from *Toxicodendron*. These cp genomes ranged from 158,221 bp (triploid *T. verniciiflum*) to 159,710 bp (*T. verniciiflum*) and comprised all the four major components of chloroplast genome architecture. The size of genome, content of GC, length of IR, LSC, and SSC regions and gene content exposed high similarity among the genomes, suggesting that *Toxicodendron* species shared low diversity [24–25, 35]. 116 genes, including 82 protein coding genes, 30 tRNA genes and four RNA genes were annotated in triploid *T. verniciiflum* cp genome. The GC content is closely related to species affinity [36]. High GC content is conducive to the stability of the genome and maintaining the complexity of the sequence. The four rRNAs genes have high GC content, which results are a high GC content in the IR regions [37]. Usually, a higher GC content indicated a more stable genome sequence [38].

The expansion and contraction of IR and SC boundaries are thought to be the main cause of CP genome size changes, although CP genomes in plants are highly conserved [39]. The change of the IR and SC junction is common phenomenon and plays an important role in evolution [40–41]. After comparing CP genomes among the five *Toxicodendron* species in this study, the results showed that the boundary region between the SC and the two IR regions was relatively conserved, with gene distribution and specific location exhibiting high consistency.

Previous studies have shown that tRNA activity may be a key factor triggering the inversions events [42]. Small inversions in the cp genome of angiosperms are ubiquitous and commonly associated with a hairpin secondary structure and are generally detected by performing pairwise comparisons between sequences of closely related taxa [26]. Many small inversions are generated by parallel or back mutation events during chloroplast genome evolution [26, 43]. In the present study, seven small inversions were discovered based on the sequence alignment of the five complete cp genomes. Two of them (ccsA-ndhD and trnS-psbZ) were specific to triploid *T. verniciiflum*, whereas inversions in *tmG-tmR* and *tmQ-psbK* occurred in *T. verniciiflum*. The inversion in *ccsA-ndhD* has been reported in other studies [16, 44–46]. These small inversion regions will provide abundant information for marker development in phylogenetic analyses of related *Toxicodendron* species [43, 47–48].

However, small inversions of noncoding sequences may influence sequence alignment and character interpretation in phylogeny reconstructions, so caution is necessary when using cp noncoding sequences for phylogenetic analysis [46, 49].

To solve phylogenetic problems at the species level, or to identify species using DNA barcodes, the high evolutionary rates regions were needed to identify [50]. Seven intergenic spacer regions including ccsA-ndhD, trnS-psbZ, atpF-atpH, trnW-trnP, rpl16 intron, trnG-trnR and trnQ-psbK are highly variable regions in the Toxicodendron chloroplast genome, which can be used as candidate DNA barcodes for future studies. These variable regions may also be useful for assessing phylogenetic relationships and interspecific differences of Toxicodendron species [51].

Complete chloroplast genome provides sufficient information sites for resolving phylogenetic relationships of plant, and have been examined to be effective in the ability of differentiation in lower taxonomic levels and provide valuable data for resolving complex evolutionary relationships [52–53]. In the current study, all Toxicodendron species were divided into main clades; one clade contained four species (T. verniciuum cv. Dahongpao, T. verniciuum cv. Hongpigaobachi, T. verniciuum cv. Yanggangdamu, and triploid T. verniciuum), and the species of T. verniciuum cv. Dahongpao, triploid T. verniciuum are natural triploid, The other clade contained T. succedaneum 1#, T. verniciuum, T. sylvestre and T. succedaneum and T. verniciuum, and all of the species are diploid [25].

Compared to morphological trait classification systems, molecular markers can reveal genetic differences at the DNA reveal genetic differences and are effective for evaluating the genetic diversity of germplasm in breeding programs [54]. SSR markers were tested to be a more advanced tool than all of these markers, which are very suitable for the identification and classification of Toxicodendron species [55]. In our study, the natural triploid samples and diploid samples can be distinguished clearly by these SSR markers. All triploid samples grouped together in group I, and all diploid accessions grouped together in group II.

Conclusions

The current study primarily explored the chloroplast genome of triploid T. verniciuum and compared it with related species within Toxicodendron genus. The size of genome, structure and organization of gene were shown to be conservative, which is similar to those reported cp genomes of Toxicodendron species. Five hotspot regions were identified and may be utilized as potential molecular markers for population genetic and phylogenetic studies in Toxicodendron. Phylogenetic analysis based on cp genomes and SSR markers both showed that triploid T. verniciuum can be distinguished with T. verniciuum cv. Dahongpao, while T. verniciuum cv. Dahongpao had a close relationship with T. verniciuum cv. Hongpigaobachi, which we speculated that T. verniciuum cv. Hongpigaobachi may be a natural triploid. Therefore, complete cp genomes is useful for species identification, taxonomic clarification, and genomic evolutionary analysis. Further research on the relationships within Toxicodendron genus and the identification of ploidy should incorporate morphology and genome wide analyses to enhance the results.

Materials And Methods

Taxon sampling and morphological analysis

T. verniciuum cv. Dahongpao (DHP) and T. verniciuum cv. Hongpigaobachi (GBC) were cultivated in the germplasm resource nursery in Yangling Shaanxi Province, triploid T. verniciuum (TZT), were cultivated in the germplasm resource nursery in Zhaotong Yunnan Province, while T. succedaneum (TRB) and T. verniciuum (TZG) were cultivated in the germplasm resource nursery in Southwest Forestry University, in Kunming Yunnan Province which were introduced from Changsha Hunan Province and Wenshan Yunnan Province, respectively. In order to compare the differences of leaf shapes among the five accessions, three leaves from each of the species were measured for five foliar traits, and all the variables can be directly measured. The five foliar traits were: (1) leaf length; (2) leaflet number; (3) leaflet length; (4) leaflet width; and (5) leaf shape index. All variables were measured using a ruler (0.1 mm resolution) and are recorded in mm. Three samples were selected from each accession, and each sample was tested for three biological replicates.

Dna Extraction

Fresh leaves of T. verniciuum cv. Dahongpao and T. verniciuum cv. Hongpigaobachi were collected from Yangling Shaanxi Province (108.08E, 34.27N), triploid T. verniciuum was collected in Zhaotong Yunnan Province (103.72E, 27.34 N), while T. succedaneum and T. verniciuum were collected from germplasm resources in Southwest Forestry University in Kunming Yunnan Province (102.76 E, 25.06 N), quality is frozen in liquid nitrogen, and stored at ultra-low-temperature refrigerator at -80°C until use. The voucher specimen deposited in the Herbarium of Southwest Forestry University. The total genomic DNA was extracted with the TGuide plant genomic DNA prep kit (Tiangen Biotech, Beijing, China) and DNA quality was inspected in 0.8% agarose gels, DNA quantification was performed using a NanoDrop spectrophotometer, DNA samples were stored at -80°C at the Key Laboratory of State Forestry Administration on Biodiversity Conservation in Southwest China, Southwest Forestry University, Kunming, China.
Genome Sequencing, Assembly And Annotation

Total DNA was used to generate libraries with an average insert size of 350 bp with the Illumina Novaseq 6000 platform. Approximately 8 Gb raw data were produced with 150 bp pair-end read lengths. The complete cp genome assembly using GetOrganelle software [56] was used to assemble the complete cp genome of the four species, with *Pistacia weinmannifolia* as the reference. Geneious R8 (Biomatters Ltd, Auckland, New Zealand) software was used for initial cp genome annotation. Start and stop codons were checked and adjusted manually when necessary by comparing them to the reference genome *P. weinmannifolia*. The tRNA genes were further confirmed through online tRNA scan-SE web servers [57]. The gene map of annotated *Toxicodendron* chloroplast genome was drawn by OGDRAW online [58]. The annotated sequences have been deposited to the NCBI GenBank database under the accession numbers OP235457, OP271782, OP279729 and OP279730 (Corresponding to *T. succedaneum*, triploid *T. verniciflum*, *T. verniciflum* cv. Hongpigaobachi and *T. verniciflum*).

Comparative Analysis Of The Chloroplast Genomes Of The Four Accessions

To investigate divergence in cp genomes, identity across the whole cp genomes was visualized using the online genome comparison tool mVISTA viewer with Shuffle-LAGAN mode among the five accessions with *T. verniciflum* cv. Dahongpao [24] as a reference to show inter-and intraspecific variations [59]. The software MAUVE alignment [60] was employed to analyze and compare the plastome structure of triploid *T. verniciflum* with the other four species of *Toxicodendron*. Furthermore, events of IR expansion and contraction were compared between these species, the junction regions between the IR, SSC, and LSC of five species were compared using the online program IR scope [61].

To identify the mutational hotspot regions, we calculated nucleotide diversity (Pi) across the whole plastome. Eight *Toxicodendron* (*T. succedaneum* 1# (MT211614), *T. sylvestre* (MT211615), *T. verniciflum* cv. Yanggangdamu (MK419151), *T. succedaneum* (OP235457), *T. verniciflum* (OP279730), triploid *T. verniciflum* (OP271782), *T. verniciflum* cv. Hongpigaobachi (OP279729), *T. verniciflum* cv. Dahongpao (MK550621)) plastome sequences were aligned using MAFFT version 7 software [62] and the nucleotide diversity was detected using DnaSP version 5 software [63] with sliding window strategy. The step size was set to 200 bp, with a 600 bp window length.

Amplification Of Ssr Marker

Genomic DNA of 15 samples from 5 species was extracted from young leaves of each sample using the TGuide plant genomic DNA prep kit (Tiangen Biotech, Beijing, China), DNA quality was inspected in 0.8% agarose gels, and DNA quantification was performed using a NanoDrop spectrophotometer. From among the 116 SSR primers [64–68], 7 primer pairs producing amplicons in all the five species. Sequence information of the SSRs is listed in Table 5. SSR analysis was carried out using 7 primer pairs, the forward of which (F) was labeled 5(6)-carboxyfluorescein (FAM) by the 5’-terminal. PCR reactions were carried out in 25 µL, containing 12.5 µL 2×PCR Mix, 1 µL of each primer, respectively, 1 µL template DNA. The PCR cycling conditions were as follows: pre-denaturation at 94 °C for 4 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at the corresponding temperature (52 °C-61 °C) for 30 s, extension at 72 °C for 1 min and the final extension for 10 min at 72 °C. Amplified SSR alleles were rechecked in 1.5% binary gels by electrophoresis at 80 volts for 50 min and gel documentation was performed with Gel Logic 200 imaging system. Capillary electrophoresis-based fragment analyses of single pollen amplified SSR alleles were conducted on an ABI3730XL Genetic Analyzer following the manufacturer’s instruction to generate GeneScan files. Scoring of alleles was done using ‘GeneMaker’ version 2.2. Genetic distance among the 15 accessions was calculated according to the software GenoDive [29] and the cluster analysis was carried out based on the matrix using MEGA 5.0.

![Table 5: Information of SSR markers](image)

<table>
<thead>
<tr>
<th>Primers</th>
<th>F(5’-3’)</th>
<th>R(5’-3’)</th>
<th>Repeat motif</th>
<th>Tm/ °C</th>
<th>Sequence length/bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>M156</td>
<td>AAGCTAGCAATACACATAGG</td>
<td>CTGACAAGTTCCAGAGCAGGG</td>
<td>(CA)14(CT)9N(AAT/C)16</td>
<td>56</td>
<td>120–152</td>
</tr>
<tr>
<td>AG28</td>
<td>TATCGCATAGGGTTCCA</td>
<td>CGGGATGGAGCCGCCAATGA</td>
<td>(GGA)15</td>
<td>61</td>
<td>222–230</td>
</tr>
<tr>
<td>M18</td>
<td>AGGCTCCAAATCCATGCTC</td>
<td>CAAGAGCAAGAACATAGAATAA</td>
<td>(AAGA)27</td>
<td>52</td>
<td>187–195</td>
</tr>
<tr>
<td>B127</td>
<td>GAAGGTGCTAACCCTTCTGATA</td>
<td>GCTATGGGGTATCTGATGTTT</td>
<td>(TC)27</td>
<td>54</td>
<td>231–257</td>
</tr>
<tr>
<td>B004</td>
<td>TCTGGTACAGGTTGCTATTGTACAGG</td>
<td>GATGATGACATCAATCCAAACAAA</td>
<td>(TC)15</td>
<td>57</td>
<td>216–252</td>
</tr>
<tr>
<td>B095</td>
<td>TGGAAGCGACAGTAACTCATAGAG</td>
<td>ACTCTTTTCTGTAAAAATTGTC</td>
<td>(TC)14</td>
<td>52</td>
<td>163–187</td>
</tr>
<tr>
<td>B041</td>
<td>ATTCCTTCCCATAAGGATCCATTC</td>
<td>TACCTAGTGAGGGAGGAAAGAGA</td>
<td>(CT)14</td>
<td>56</td>
<td>174–288</td>
</tr>
</tbody>
</table>
Phylogenetic Analysis

Representative samples of species were chosen from complete, annotated plastid genomes of *Toxicodendron* to investigate plastid genomic structure variation across the order. This included eight completely sequenced and annotated plastomes (4 from NCBI GenBank [*T. verniciflum* cv. Dahongpao (MK550621), *T. succedaneum* cv. 1# (MT211614), *T. sylvestre* (MT211615) and *T. verniciflum* cv. Yanggangdamu (MK419151)]) plus the four newly sequenced [*T. verniciflum* cv. Hongpigaobachi, triploid *T. verniciflum*, *T. succedaneum* and *T. verniciflum*] *Toxicodendron* plastomes), and two out-group: *Pistacia weinmanniifolia* (MF630953) and *Mangifera indica* (KY635882). The nucleotide sequences were aligned using MAFFT version 7 software [62]. The phylogenetic analyses were performed using maximum likelihood (ML) and Bayesian inference (BI). ModelFinder [69] was used to select the best-fit model with default setting and the maximum likelihood (ML) analysis was performed using IQ-TREE 1.5.5[70] with 1000 bootstrap replications. The BI analysis was performed by Mybayes 3.2.6 [71]. The jModelTest 2.0 program [72] was used to determine the best-fitting model for each dataset based on the Akaike information criterion and the optimal model of “GTR + F + I + G4”. The Markov chain Monte Carlo (MCMC) algorithm was run for 1,000,000 generations, and a burn-in of 25% was used for the analysis.

Declarations

Authors’ contributions

Dan Zong designed the experiments and organized the manuscript; Zhengsheng Qiao, Jintao Zhou, Peiling Li, Peihua Gan, Meirong Ren analyzed and interpreted the data.

Dan Zong and Chengzhong He Original edited the manuscript;

All authors revised, read, and approved the final version of the manuscript.

Acknowledgements

Not applicable.

Funding

This work was supported by Yunnan Provincial Expert Workstation (202005AF150020) and Yunnan Provincial Ten-Thousand Program - Industry Leading Talents (YNWR-CYJS-81210420) and Applied Basic Research Foundation of Yunnan Province (202101AU070144).

Availability of data and materials

The datasets generated during the current study are available in the NCBI GenBank (Accession number OP235457, OP271782, OP279729, OP279730)

Ethics approval and consent to participate

We have obtained the permission of Southwest Forestry University to collect the species of *Toxicodendron*. The collection and usage of plant specimens in current study complied with the IUCN Policy Statement on Research Involving Species at Risk of Extinction and is allowed by the Convention on the Trade in Endangered Species of Wild Fauna and Flora. Ethical approval was not applicable for this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1Key Laboratory for Forestry Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming, China. 2Key Laboratory for Forest Genetics and Tree Improvement & Propagation in Universities of Yunnan Province, Southwest Forestry University, Kunming, China. 3Key Laboratory of Biodiversity Conservation in Southwest China, State Forestry Administration, Southwest Forestry University, Kunming, 650224, China.
References


Page 11/18


Figure 1

Gene map of the five Toxicodendron cp genomes. Genes shown outside the outer circle are transcribed clockwise and those inside are transcribed counterclockwise. Genes belonging to different functional groups are color-coded. The dashed area in the inner circle indicates the GC content of the cp genome of Toxicodendron.
Figure 2

Comparison of the LSC, IR, SSC junction positions among five *Toxicodendron* cp genomes.
Figure 3

Identity plots comparing the cp genomes of five *Toxicodendron* accessions using *T. vernicifluum* cv. Dahongpao as a reference sequence. The vertical scale indicates the percentage of identity, ranging from 50 to 100%. The horizontal axis indicates the coordinates within the cp genome. Genome regions are color coded as protein-coding, rRNA, tRNA, intron, and conserved non-coding sequences (CNS).
Figure 4

MAUVE alignment of five Toxicodendron accessions cp genomes. The *T. vernicifluum* cv. Dahongpao genome is shown at the top as the reference genome. Within each of the alignments, local collinear blocks are represented by blocks of the same color connected by lines.

Figure 5

Sliding window analysis of the whole cp genomes of eight Toxicodendron accessions. Window length: 600 bp, step size: 200 bp. X-axis, the position of the midpoint of a window; Y-axis, nucleotide diversity of each window.
Figure 6

Predicted hairpin loops of small inversions in the five plastomes of *Toxicodendron*.

Figure 7

Phylogeny of eight *Toxicodendron* accessions inferred from ML and BI analyses of different cp genome sequences.
Figure 8

Dendrogram generated by UPGMA from genetic distance data based on SSR markers

Supplementary Files

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

- SupplementaryMaterial.zip