The chromosome-level genome for *Toxicodendron vernicifluum* provides crucial insights into Anacardiaceae evolution and urushiol biosynthesis.

**Highlights**

- We provide the first chromosome-level genome for *T. vernicifluum*
- We explore the phylogenetic position of lacquer tree
- We identify genes involved in the urushiol and lignin biosynthetic pathways
- Analysis of the unique active metabolites in cultivar lacquers by LC-MS/MS is reported

Guoqing Bai, Chen Chen, Chenxi Zhao, ..., Xiaofeng Cong, Yun Jia, Sifeng Li
lisf60@sina.com
The chromosome-level genome for *Toxicodendron vernicifluum* provides crucial insights into Anacardiaceae evolution and urushiol biosynthesis

Guoqing Bai,1,7 Chen Chen,1,7 Chenxi Zhao,2,6,7 Tao Zhou,3,7 Dan Li,4,7 Tianhua Zhou,5 Weimin Li,1 Yuan Lu,1 Xiaofeng Cong,1 Yun Jia,1 and Sifeng Li1,8,*

**SUMMARY**

The lacquer tree (*Toxicodendron vernicifluum* (Stokes) F.A. Barkley) is an important tree with economic, industrial, and medicinal values. Here, we generated the reference genome of *T. vernicifluum* at the chromosome level with 491.93 Mb in size, in which 98.26% of the assembled contigs were anchored onto 15 pseudo-chromosomes with the scaffold N50 of 32.97 Mb. Comparative genomic analysis revealed the gene families related to urushiol biosynthesis were expanded, contributing to the ecological fitness and biological adaptability of the lacquer tree. We combined multi-omics data to identify genes that encode key enzymes in the *T. vernicifluum* urushiol and lignin biosynthetic pathways. Furthermore, the unique active metabolites, such as butin and fisetin, in cultivar lacquers were identified by metabolism profiling. Our work would provide crucial insights into metabolite synthesis such as urushiol and lignin, meanwhile offer a basis for further exploration of the cultivation and breeding of *T. vernicifluum* and other Anacardiaceae members.

**INTRODUCTION**

*Toxicodendron vernicifluum* (Stokes) F. A. Barkley belongs to Anacardiaceae family, also known as the Chinese lacquer tree (Wang et al., 2020a). Like rubber trees, the lacquer tree can produce a sap called “raw lacquer” by cutting its phloem, which is an excellent adhesive and paint material with multiple properties such as anticorrosion, antrust, non-oxidation, and resistance to acid, alcohol, and high temperature. The lacquer tree is native to China and the Indian subcontinent, and it has been farmed for thousands of years in East Asian nations, such as Korea and Japan (Fu et al., 2007; Nie et al., 2009). Historically, raw lacquer was traditionally produced of lacquerwares that were used as tableware and sacrificial utensils in Asian cultures (Li et al., 2016; Ma et al., 2015). Currently, Chinese raw lacquer manufacturing accounts for 85% of the global total output value, and plentiful lacquer tree germplasm resources were available in China, such as “Da-hongpao”, “Gaobachi”, “Huangmao Guizhou”, and “Hongmao Guizhou” etc.

The most common product obtained from the lacquer tree is raw lacquer, which has a wide range of applications in aesthetic decoration, petroleum, nautical, and other industrial fields (Lu et al., 2014; Xu et al., 2012; Yang et al., 2015). Raw lacquer has been shown to be a natural polymer made up of various components, including urushiol, laccol, thitsiol, laccase, polysaccharides, and glycoprotein (Honda et al., 2008; Yang et al., 2018). Urushiol is the fundamental component for quick drying and film formation of raw lacquer, and it has a long history as a natural performance. Many novel coating materials based on the structure of urushiol have been created recently (Watanabe et al., 2016; Xue et al., 2020). Urushiol, in particular, has been shown to represent a collection of closely related chemicals that can cause contact dermatitis. (Bonnekoh et al., 2018; Lianchezhian et al., 2012; Rojas-Muñoz et al., 2012). Nonetheless, Urushiol has been studied for their pharmacological properties, which include anticancer, antibacterial, and antioxidant properties. (Cho et al., 2015; Kim et al., 1997; Xie et al., 2016; Zhao et al., 2009).

Urushiol is a mixture of phenols (catechol and resorcinol) with a lengthy side chain of 15–17C and contains multiple terminal vinyl structures (Kim et al., 2019a, 2019b). It has been shown that fatty acid metabolism intermediates, such as hexadecanoyl-CoA, alkyl tetraketone are significant in the formation of urushiol (Bai et al., 2018; Weisberg, 2014). The phenylpropanoid metabolic pathway might also produce catechol
and resorcinol. According to transcriptome and enzyme functional investigations, type III polyketide synthase (PKS) was thought to be the first enzyme to catalyze urushiol production, with OXSM/FabF, FabZ, FATB, and ACSL all taking part. However, the biosynthesis routes of urushiol remain largely uncharacterized.

Thanks to the emergence and advancement of high-throughput sequencing technology, almost 600 complete plant genome assemblies are now accessible in public sources (Kersey, 2019). The high-quality genomes of commercial plants in Anacardiaceae, such as the “king of fruits” Mango, and the “important nut crops” Pistachio, have been sequenced recently (Wang et al., 2020b; Zeng et al., 2019). These genomes provide a foundation for studying genomic-based trait development and specialized Anacardiaceae biochemistry underneath. Furthermore, whole-genome duplication (WGD) or polyploidization is major drivers of specialization and the creation of new features and functions. These duplicates make them available for the evolution of a range of activities, such as producing insect resistance, boosting tolerance to stress, and encouraging plants to acquire adaptive benefits, all of which contribute to improved plant fitness (Jiang et al., 2021; Kang et al., 2020; Ma et al., 2021; Tang et al., 2016). Terpenoid biosynthesis genes, for example, are clustered and co-expressed, lowering the risk of incomplete transcription in Lamiaceae due to the loss of a single gene participating in the process (Consortium, 2018; Li et al., 2021; Xu et al., 2016; Zhao et al., 2019). Furthermore, the mango genome’s increased cluster of chalcone synthase genes (CHS) has co-synteny expression, which is linked to specific phenolic compounds (Wang et al., 2020b). Those specific biochemistries are extremely important for plant evolution and fitness.

Despite the considerable importance of the lacquer tree, genetic information on the species is scarce, which has hindered its study and utilization. Here, we reported the chromosome level of the lacquer genome, which is the first sequenced species in the genus Toxicodendron. Using genomic, transcriptomic, and metabolomic techniques, we thoroughly screen and discover the candidate genes responsible for urushiol and lignin production in the lacquer tree. The availability of our genomic data will be helpful for exploration of urushiol biosynthesis and molecular breeding and engineering of lacquer trees.

RESULTS
Chromosome-level assembly and annotation of T. vernicifluum genome
A total of 130.8 Gb of short reads and 106.28 Gb of ONT long reads were obtained based on DNBSEQ and the Oxford Nanopore platform, respectively. After quality control and trimming, there were 124.4 Gb of DNBSEQ clean reads and 105.28 Gb of ONT high-quality reads. Based on 17-mer frequency distribution analysis with short reads, the T. vernicifluum (Figure 1A) genome was estimated in size of 560.06 Mb (Figure S1, Tables S1) with relatively high heterozygosity (0.56%) and repetitive sequence content (69.02%) (Table S1). The predicted genome size based on 17-mer analysis was close to the genome size (~540 Mb) determined experimentally by flow cytometry (Figure S2).

After assembling, we obtained a 491.81 Mb assembly with a contig N50 length of 5.26 Mb by using SMART denovo and polished it by racon, medaka, and pilon (Figure 1D and Table 1). The Benchmarking Universal Single-Copy Orthologs (BUSCO) evaluation score was used to assess the completeness and quality of the genome assembly, which resulted in 98.9% of the single-copy orthologs’ completeness (Table S3), indicating a very complete and high-quality genome assembly. To assist the chromosome-level assembly, 130.74 x Hi-C data were generated on the DNBSEQ platform. Finally, 491.93 Mb super-scaffolds with N50 of 32.97 Mb was generated based on ~73.22 Gb of high-throughput chromatin conformation capture (Hi-C) data using the hierarchical clustering strategy. In total, 98.26% of the assembled sequence was anchored and oriented onto 15 pseudochromosomes with lengths ranging from 25.95 to 42.48 Mb in total (Table S4). The genome of T. vernicifluum had a GC content of 35.7% (Figure S4 and Table S5).

The full-length transcriptome of a mixed library of leaves and phloem was generated and sequenced using the PacBio platform to aid with genome annotation (Table S6). We identified 303.34 Mb (61.66% of the genome length) of repetitive sequences in the genome assemblies of T. vernicifluum based on a combination of homology-based and de novo approaches (Table S7). Among them, 266.81 Mb of long terminal repeat (LTR) retrotransposons accounted for the largest proportion and made up 54.24% of the genome (Table S8, Figure S5). Totally, we predicted 32,682 protein-coding genes in the genome of T. vernicifluum with an average length of 3993.57 bp and an average exon number of 10.17 for each gene (Table 1) based on the homology annotation, de novo prediction, transcript mapping, and further
filtering. 96.16% (31,428) of the protein-coding genes in the *T. vernicifluum* genome were assigned to multiple functions based on the analysis via various protein database, including NR, SwissProt, Kyoto Encyclopedia of Genes and Genomes (KEGG), KOG, TrEMBL, and Inter-Pro. The gene distribution and GC content along each chromosome were calculated and showed in Figure 1B. In addition, the *T. vernicifluum* genome contained 104 miRNAs, 510 tRNAs, 498 rRNAs, and 4,096 small nuclear RNAs (snRNA) (Figure 1B and Tables 1 and S9).

**Phylogenomics and divergence time estimation**

To better comprehend the phylogenetic placements of Anacardiaceae among angiosperms, 1,914 single-copy genes shared by *T. vernicifluum* and other 14 plant taxa were recovered for the phylogenomic interference, with *Oryza sativa* was set as an outgroup. The phylogeny confirmed the monophyly of malvids. All Anacardiaceae, Rutaceae, and Sapindaceae species were discovered to be clustered in the same clade, showing close relationships of these three families. In particular, all Anacardiaceae species (*Anacardium occidentale*, *T. vernicifluum*, *Mangifera indica*, *Sclerocarya birrea*, and *Pistacia vera*) were clustered into one monophyletic group (Figure 2A) with a high bootstrap support value, and *T. vernicifluum* was closely related to *P. vera*. The phylogenetic tree indicated that the malvid clade diverged ~95.5 million years ago (Mya) (88.9–101.4 Mya). The splits in the crown group Sapindales were estimated to have occurred 74.3 million years ago (Mya) (67.7–81.7 Mya), while the Anacardiaceae divergence time was estimated to be 52.3 Mya (42.5–61.9 Mya). The divergence between the Trib. Rhoideae clade and the Trib. Anacardieae clade occurred 29.6 Mya (24.7–35.4 Mya) for four species, and it was also demonstrated that *T. vernicifluum* and *P. vera* belonging to Trib. Rhoideae diverged 25.2 Mya (13.5–27.8 Mya).

**WGD and collinearity analyses**

Here, the density distribution of synonymous substitution rates per gene (Ks) between collinear paralogous genes was utilized to detect the WGD events. After calculating the Ks values of duplicate gene pairs, we found the values of Ks values for the collinear gene pairs peaked at 1.8, corresponding to an ancient WGD event that occurred around 47.3 Mya. We found that the peak value of orthologs between *T. vernicifluum* and *P. vera* (Ks = 0.1) was lower than the value of Ks = 0.4 between *T. vernicifluum* and *S. birrea*, indicating that the speciation between *T. vernicifluum* and *P. vera* occurred later, and this result

**Figure 1. Genomic landscape of lacquer tree**

(A) Tree, leaf, flower, and raw lacquer blooding of *Toxicodendron vernicifluum*. (B) Overview of lacquer genome assembly. (I) Chromosomes of *T. vernicifluum*; (II) density of the genome (500-kb sliding windows); (III) TE density of the genome (500-kb sliding windows); (IV) non-coding RNA content of the genome (red lines represent miRNA, blue lines represent tRNA, green lines represent snRNA, purple lines represent rRNA); (V) paralogous genes on different chromosomes.
was supported by the phylogeny. Based on the distribution of Ks, we found that *T. vernicifluum* may have no recent specific WGD, and it probably shared a WGD event with other Anacardiaceae species before the diversification of Anacardiaceae. The sharp peaks in 4-fold synonymous third codon transversion (4DTV) values ($K_s = 0.5$) also suggest the *T. vernicifluum* genome shared a WGD event with *M. indica* (Figures 2B and 2C).

Synteny block analysis is often used to determine chromosome evolution among related species. Here, we analyzed the aligned protein sequences of *T. vernicifluum* in comparison with its close relatives *M. indica*. A total of 19,987 syntentic blocks, containing 210,759 pairs of collinear genes, were identified in the *T. vernicifluum* genome, and there were 64,444 collinear gene pairs from 4,867 collinear blocks detected between *T. vernicifluum* and *M. indica*. The comparative genome structure between *T. vernicifluum* and *M. indica* showed relatively high collinearity (Figure 2D). For several collinear regions, the chromosomes of *M. indica* showed a one-to-one syntenic relationship with *T. vernicifluum*. For instance, *MiChr8*, *MiChr15*, *MiChr20* of *M. indica* corresponded to *TvChr2*, *TvChr5*, *TvChr10*, and *TvChr8*. Interestingly, *TvChr4* of *T. vernicifluum* corresponded to *MiChr13* and of *MiChr17* *M. indica* (Figure 2D), indicating that these two chromosomes of *M. indica* might be derived from the chromosome duplication. In addition, we also found that most chromosomes of *T. vernicifluum* matched to more than one chromosome of *M. indica*. For instance, *TvChr1* corresponded to *MiChr1*, *MiChr10*, *MiChr12*, and *MiChr18*; *TvChr2* corresponded to *MiChr1*, *MiChr3*, *MiChr4*, *MiChr5*, *MiChr7*, *MiChr8*, *MiChr11*, *MiChr14*, *MiChr16*, *MiChr19*, and *MiChr20*; *TvChr3* corresponded to *MiChr1*, *MiChr6*, *MiChr10*, and *MiChr12*. Therefore, it can be inferred that most chromosomes in *M. indica* might have formed by fragmentation and recombination after the divergence between *T. vernicifluum* and *M. indica*, and our analyses also indicated that there was a recent WGD for *M. indica*, but not for *T. vernicifluum*.

### Expansion and contraction of gene families for *T. vernicifluum*

As a result, a total of 15,917 gene families comprising 27,163 genes were homologously identified in the lacquer genome (Table S11), which included 9,945 gene families shared among the four species (*A. occidentale*, *M. indica*, *P. vera*, and *S. birrea*), and 539 gene families containing 1,559 genes were specific to *T. vernicifluum* (Figure 2E). GO enrichment analysis of unique gene families revealed that these

| Parameter                          | *T. vernicifluum* |
|------------------------------------|-------------------|
| Estimate of genome size (Survey)   | 560.06 M          |
| Estimate of genome size (Hi-C)     | 491.93 M          |
| Total length of contigs            | 491.81 M          |
| Total number of contigs            | 866               |
| N50 of contigs (bp)                | 5.26 M            |
| Largest contg (bp)                 | 19.36 M           |
| Total length of super-scaffolds    | 491.93 M          |
| N50 of super-scaffolds             | 32.97 M           |
| GC content                         | 35.77%            |
| Complete BUSCOs %                  | 98.9              |

**Table 1. Summary of genome assembly and annotation of *T. vernicifluum* genome**

---

4 iScience 25, 104512, July 15, 2022
specific gene families were considerably enriched in cysteine-type peptidase activity (GO: 0008234), peptidase activity (GO: 0070011), defense response (GO: 0006950), and response to stress (GO: 0006950) (Table S12). A KEGG pathway analysis of *T. vernicifluum* specific gene families revealed significantly enrichment in genes involved in endocytosis, metabolic pathways, phenylpropanoid biosynthesis, and plant-pathogen interaction, possibly contributing to the ecological fitness and biological adaptability of the lacquer tree (Table S13, Figure S8).

To further explore the significant expansion and contraction of gene families in the evolution process, an OrthoFinder analysis of *T. vernicifluum* and other 14 species was performed. Our results showed that 438 families exhibited significant expansion, while 1,707 gene families exhibited significant contraction (Figure 2A). KEGG pathway analysis of the expanded gene families showed marked enrichment in genes involved in functions, such as phenylpropanoid biosynthesis, biosynthesis of unsaturated fatty acids, metabolism of terpenoids and polyketides, and plant-pathogen interaction (Table S14, Figure S9). And the GO enrichment analysis revealed that the expanded gene families were enriched in binding, defense, oxygen oxidoreductase activity, and lignin catabolic process (Table S15). On the other hand, the contracted gene families were enriched in functions related to the phosphotransferase activity, transporter activity, and catalytic activity (Table S16). Furthermore, the contracted gene families showed the marked enrichment in...
Determination of genes involved in urushiol biosynthesis and a candidate polyketide synthase family

The urushiol synthesis pathway is illustrated in Figure 3A, where the enzymes and their orthologs found in T. vernicifluum genome are indicated according to the expression profile of genes related to urushiol biosynthesis. In total, we found 79 orthologs participating in urushiol biosynthesis process (Tables S18 and S19). There are enzymes associated with the synthesis of long-chain fatty acid and their derivatives, including long-chain acyl-CoA synthetase (ACSL), long-chain 3-ketoacyl-CoA synthase (KCS), acetyl-CoA carboxylase biotin carboxyl carrier protein (accB), fatty acyl-ACP thioesterase A (FAT), fatty acyl-ACP thioesterase B (FATB), and 3-oxoacyl-[acyl-carrier-protein] synthase II (FabF), as well as enzymes related to polyketide production, such as polyketide synthase (PKS), reductases such as 3-hydroxyacyl-[acyl-carrier-protein] dehydratase (FabZ), and laccase (LAC). Moreover, RNA-seq was also used to screen the expression levels of urushiol-biosynthesis-related genes from four cultivars and a wild lacquer tree. As a result, we discovered that seven urushiol-biosynthesis-related genes, including KCS11, KCS4, and TvPKS9, 16, 19, 22, 23, were expressed lower in the wild type while they were expressed higher in the cultivars (Figure S11). Our metabolite profile in the five varieties of lacquer trees supports this conclusion. The results showed that urushiol contents, including 3-[8Z,12'E,15'Z-pentadecatrienyl]-catechol, 3-[8Z,11'Z,14'-pentadecatrienyl]-catechol, 3-[9E,11E,13Z]-pentadecatrienyl-1,2-catechol and 3-[8,12-heptadecadienyl]-1,2-catechol, 3-[10Z,13E]-10,13-pentadecadienyl-1,2-catechol, urushiol III, and 3-[8Z,11E]-8,11-pentadecadienyl-1,2-catechol, were significantly greater in cultivars than in wild lacquer tree (Table S21).
PKS played a vital role in the initial phase of urushiol biosynthesis and facilitated to yield the key intermediate. We identified 33 orthologs of PKS clustered in Chr 2, Chr 6–8, Chr 12, and Chr 13 using the conserved domain searched in T. vernicifluum genome (Table S19). Tve06G1275, Tve06G1276, Tve06G1281, Tve06G1282, Tve06G1283, Tve06G1288, and Tve06G1292 were closely connected to M03G0182300 on Chr 3 of the M. indica genome. Similarly, the clusters of Tve13G0944, Tve13G0947, Tve13G0997, Tve13G0999, Tve13G1000, and Tve13G1003 were very closely related to Mi19G0008700 on Chr 9 of M. indica (Figure 3B). Besides, to evaluate the evolutionary relationships of 33 PKS genes derived from T. vernicifluum, all the protein sequences were aligned to construct an unrooted ML phylogenetic tree. PKS genes were found to be divided into three distinct subgroups, implying that their structure and function are divergent (Figure 3C). To estimate who interacted with PKS proteins, we analyzed the PKS’s protein interaction network. It showed that the gene pairs formed as TvPKS13 and TvPKS14, TvPKS23 and TvPKS7, and TvPKS9 and TvPKS29 interacted with phenylpropanoid pathway enzymes, such as reductase (Tv09G1323), CHS (Tv02G03096), chalcone isomerase (Tv01G2833), and 4CL (Tv09G1674, Tv14G0514), which are involved in phenylpropanoid pathways (Figure 3D). Quantitative real-time PCR(qRT-PCR) analysis was conducted to validate the expression levels of enzyme genes involved in urushiol pathway in different cultivars, such as ACSL1, KCS3, KCS6, PKS4, and PKS11, and compared with their abundance from sequencing data. The qRT-PCR analysis showed that ACSL1 was highly expressed in GBC and HoM, and KCS2 has significantly high expression in DHP, HoM, and HuM, which is generally agreed well with the sequencing result (Figure S13).

Screening for genes encoding enzymes of the lignin biosynthetic pathway
Clarifying lacquer’s lignin production route is highly beneficial in understanding the growth law and the distinctions between natural tree and artificially produced types. Using the Basic Local Alignment Search Tool (BLAST), we identified 177 orthologous genes involved in lignin biosynthesis based on our lacquer’s genome database (Table S20), including phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate: CoA ligase (4CL), hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase (HCT), cinnamyl-CoA reductase (CCR), p-coumarate 3-hydroxylase (C3H), cinnamyl alcohol dehydrogenase (CAD), caffeoyl-CoA O-methyltransferase (CCoAOMT), caffeic acid O-methyltransferase (COMT), fumarate 5-hydroxylase (FSH), caffeoylshikimate esterase (CSE), lignin-forming anionic peroxidase (POX), and laccases (LAC). Among them, CCR catalyzes the coumarin production, plays an important role in regulating the carbon flow of lignin biosynthesis pathway. In total, we found 31 CCR homology genes in lacquer genome, while 31 are present in pear (Pyrus bretschneideri) (Cheng et al., 2017), 14 are present in soybean (Cui et al., 2021), and six are present in mulberry (Chao et al., 2021). The TvCCR phylogenetic tree constructed with CCRs from Arabidopsis thaliana, O. sativa, Populus, as well as P. vera and S. birrea revealed that CCRs are clustered into four groups (Figure S12), which indicated that CCRs had structural and functional divergence. Additionally, our transcriptome data revealed diverse expression of lignin biosynthesis genes in both wild and cultivars of lacquer tress (Figure 4), which will provide fundamental insights into engineering of lacquer tress for lignin production. Similarly, qRT-PCR analysis was carried out to analyze the expression level of four genes related to lignin biosynthesis pathway to confirm the accuracy of the Illumina expression profiles. It showed that CCR4, CCR5, and LAC5 were significantly highly expressed in GBC, and LAC5 was expressed higher in HoM, which is agreed well with the sequencing data (Figure S13).

Metabolite profiling for characterization of bioactive chemicals in lacquer trees
The extractions from lacquer tree are also used as a traditional medicinal herb in China for its various flavonoids, unsaturated fatty acids, and polysaccharides. Notably, comparative genomics revealed that 295 gene families were expanded in the category of “Biosynthesis of secondary metabolites”, and 10 for “Biosynthesis of unsaturated fatty acids”, 23 for “Flavonoid biosynthesis”, and 31 for “Metabolism of terpenoids and polyketides”, respectively (Figure S9). These enriched genes may contribute to enhancing the biosynthesis of the important bioactive metabolites such as flavonoids, unsaturated fatty acids, and urushiol biosynthesis in the lacquer tree.

The phloem of these five samples was collected for metabolite profiling by UPLC-MS/MS to examine the difference of lacquer tree cultivars “Dahongpao (DHP)”, “Hongmao Guizhou (HoM)”, “Huangmao Guizhou (HuM)”, “Gaobachi (GBC)”, and wild type (TV) in metabolism levels (Figures 5A and 5B). In the results, 12 classifications containing 862 metabolites were detected in lacquer tree phloem, including flavonoids (178), phenolic acids (166), lipids (100), organic acids (82), lignans and coumarins (31), etc (Table S21). We found differences in metabolites between cultivars and wild trees. There are 165 metabolites including flavonoid, flavanols, lignans, tannin, alkaloids, and intermediates or derivatives in urushiol biosynthesis,
such as 3-(10E)-10-pentadecenyl-1,2-benzenediol, 3-(10E)-10-pentadecenyl-1,2-benzenediol, 3-(10,13,16-heptadecatrienyl)-(Z,Z)-1,2-benzenediol, 3-(8,10,12-heptadecatrienyl)-1,2-benzenediol, 3-(8,11,14-heptadecatrienyl)-

Compared to wild lacquer tree, 231 metabolites including flavonoids, phenolic acids, lipids, organic acids, alkaloids, lignans, coumarins, amino acids and derivatives, and terpenoids were significantly changed in cultivar DHP. Similarly, there were 218 special metabolites in HoM and 287 significantly different metabolites in HuM compared with wild lacquer tree. Most phenolic acids, alkaloids, lipids, organic acids, and terpenoids were found to be abundant in cultivars than in the wild. Notably, 1-O-p-coumaroylquinic acid, cirsimaritin, eupatorin, persicoside, limocitrin-3-O-sophoroside, limocitrin-3,7-di-O-glucoside, amentoflavone-7”"O-glucoside, quercetin-3-O-2”-feruloyl) sophoroside, syringaresinol-4’”-O-(6”-acetyl) glucoside, and jasmonic acid were exclusively detected in the cultivars and not in the wild lacquer tree. Accordingly, there is also an apparent deficiency of metabolites in the cultivars. For example, we did not detect sexangularetin-3-O-glucoside-7”-O-rhamnoside in HuM, and we did not detect protocatechuic acid ethyl ester in HoM, which existed in the wild lacquer tree (Table S21). This metabolic profile among wild and cultivar lacquer trees will provide data for excellent variety breeding and comprehensive utilization of lacquer trees.

**DISCUSSION**

Lacquer genome is the first one at chromosome level in **Toxicodendron**

A high-quality reference genome for lacquer tree can facilitate lacquer molecular breeding and evolutionary research of Anacardiaceae, and it is also valuable resource for studying the molecular basis of specialized metabolites diversity and how their biosynthetic pathways evolved. Here, we generated the
The first reference genome of the lacquer tree with 491.93 Mb in size at chromosome level. The contig N50 and scaffold N50 sizes are 5.26 and 32.97 Mb, respectively, and 98.9% of the single-copy orthologs based on the BUSCO analysis, which suggested the *T. vernicifluum* genome has good assembly and the largest genome completeness (Table 1). In addition, the heterozygosity of the lacquer tree (0.56%) is lower than that of its related species, which is 0.68% and 1.72% in *M. indica* and *P. vera*, respectively (Wang et al., 2020b; Zeng et al., 2019). Owing to the lack of a genome background, the advancement of the exploration on urushiol biosynthesis pathway, gene editing, and molecular breeding of lacquer tree were hindered. In this study, the chromosome-level genome of lacquer tree not only provides a critical resource for *T. vernicifluum* functional genomic studies, particularly the investigation of the urushiol biosynthesis pathway, but also contributes to the understanding of *T. vernicifluum* biology and evolution.

**Evolution of lacquer genome**

The monophyly of malvids and close phylogenetic relationships of three Sapindales order families (Anacardiaceae, Rutaceae, and Sapindaceae) are well supported by phylogenomic analysis of 1,914 single-copy genes from 14 typical seed plant genomes. The phylogenetic signals generated in this work are consistent with the plastid genome phylogeny (Li et al., 2019). Besides, within the Anacardiaceae, we found that *T. vernicifluum* was more closely related to *P. vera*. This result is in agreement with the phylogeny based on nuclear DNA and plastid markers (Jiang et al., 2019; Wang et al., 2020a). Therefore, it means that phylogenomics is effective to rebuilt the robust phylogeny for angiosperm or even for specific taxa. However,
future phylogenomic studies should recover sufficient lineage sampling and gene sampling as well as comprehensive analytical methods to provide more convincing phylogenetic evidence for research taxa. WGD event or polyploidization seems to have occurred throughout the routes of most plant evolution, which can facilitate the emergence and specialization of novel traits and functions (Jiao et al., 2011; Wolfe, 2001). The genome of T. vernicifluum obtained in this study improved our understanding of the timing of the WGD event in Anacardiaceae. Previous studies indicated that there was no lineage-specific whole genome duplication occurred in P. vera (Guo et al., 2017), and our results also verified no lineage-specific WGD event for T. vernicifluum which might occur after the split of P. vera. The WGD time of T. vernicifluum genome (~47.3 Ma) occurred before the split of Trib. Anacardiaceae (29.6 Mya). In addition, 4DTv analyses also confirmed no specific WGD event in the T. vernicifluum. It is worth mentioning that M. indica genome shared high collinearity with T. vernicifluum, and it shared ancient WGD with other species belong to Trib. Anacardiaceae (Wang et al., 2020b). Unexpectedly, M. indica genome shared high collinearity with T. vernicifluum experienced a recent specific recent WGD (~33 MYA) which after the ancient WGD event time for T. vernicifluum (Wang et al., 2020b), and it had been found that some retained genes from WGD of M. indica were involved in the biosynthesis of secondary metabolisms. It is appealing to hypothesize that the WGD in T. vernicifluum genome may also impact the different metabolism categories with prominent contributions.

Draft maps of urushiol and lignin biosynthesis pathway
Genome sequencing is a powerful tool for studying various aspects of physiology and genetics in non-model plants, particularly in medicinal herbs and economical important species (Li et al., 2021; Ma et al., 2021; Tang et al., 2016; Zeng et al., 2019; Zhao et al., 2019). For example, waterlily, Magnolia officinalis, and persimmon draft genomes provides information into their evolutionary process (Yin et al., 2021; Zhang et al., 2020; Zhu et al., 2019), and Cucumis and Vitis genus genomes provide an understanding of important agronomic traits and stress response (Patel et al., 2020; Wang et al., 2021; Yu et al., 2021). Furthermore, many reference genomes, such as Scutellaria baicalensis, Panax Notoginseng, Senna tora, Lithospermum erythrorhizon, and garlic (Allium sativum) shed light on genes involved in the biosynthesis of specific flavonoids, terpenes, alkaloids, lignan, and a number of other secondary metabolites (Auber et al., 2020; Fan et al., 2020; Kang et al., 2020; Sun et al., 2020; Yin et al., 2021).

Phenolic compounds such as flavonoids, coumarins, and phenolic acids and their derivatives exert ecological functions as protectors against the attack of herbivores and microorganisms, attractors of specific branch of the lignin biosynthesis pathway, where it converts feruloyl-CoA to coniferaldehyde (Leple et al., 2002). CCR in plant differs structurally and functionally. For example, AtCCR1 is expressed in the flowers, leaves, and stems, undergoing a lignification process, whereas AtCCR2 is lower expressed in normal stages but highly triggered by pathogen invasion (Lauvergeat et al., 2001). CCR genes in poplar have similar properties, such as selective expression in leaves, bark, and xylem (Barakat et al., 2011). Our finding revealed that the 30 TvCCRs were clustered into four groups, whereas the most CCRs were divided into two subgroups separately (Figure S12), which indicates that TvCCRs could exist with structural and functional divergence from CCRs in other plants. In addition, in a combined transcriptome and metabolic profile, the mediates, such as cinnamic acid, p-coumaryl alcohol, and caffeic acid, along with the related genes, such as 4CL, CAD, were significantly expressed higher in cultivars than in the wild, implying that artificial selection or cultivation has altered lacquer lignin synthesis and improved its
woody properties. The lignin biosynthesis pathway of the lacquer tree has been identified, laying the groundwork for future molecular breeding and cultivation studies.

Urushiols are produced in many Anacardiaceae plants, particularly those in the genus Toxicodendron, such as *T. vernicifluum* (lacquer tree), *T. radicans* (poison ivy), and *T. diversilobum* (poison oak), and often cause contact dermatitis in sensitive individuals (Mina et al., 2017). Urushiol is alk-(en)-yl catechol derivatives with a 15- or 17-carbon side chain that are probable insect, fungus, or herbivorous defense compounds (Kim et al., 2019a, 2019b). Urushiol has been shown to have substantial antioxidant, immune-enhancing, neuroprotective, anti-inflammatory, and antitumor properties in numerous studies (Cho et al., 2012; Kim et al., 2015c, 2015d; Lee et al., 2004, 2015, 2017). Kadokura et al. also found two novel urushiols with HIV-1 reverse transcriptase inhibitory activity, indicating that urushiol could be a potential bioactive agent in AIDS treatment (Kadokura et al., 2015). Since 1982, many analytical chemistry techniques have been utilized successively to isolate and identify urushiol from *Toxicodendron* plants, such as gas chromatography-mass spectrometry (GC-MS), LC-MS-MS, matrix-assisted laser desorption/ionization-mass spectrometry imaging (MALDI-MSI), two-dimensional (2D) 1H-13C heteronuclear multiple quantum coherence (HMQC), and heteronuclear multiple bond coherence (HMBC) NMR spectroscopic techniques (Draper et al., 2002; ElSohly et al., 1982; Harigaya et al., 2007; Mina et al., 2017). However, little is currently known about the biosynthetic pathway of urushiol in the *Toxicodendron* genus. The urushiol biosynthesis pathway has been proposed to be involved in fatty-acid metabolism with polyketide synthase being the first step (Fu et al., 2007). Wang et al. proposed that the early genes in phenolic biosynthesis pathways like chalcone synthase (CHS) have extensive expansion in the mango genome, highlighting the importance to the metabolic flow of phenols (Wang et al., 2020b). So far, there is still a lack of details and more genes involved in this synthetic framework. In our data, we found 79 orthologs containing 33 TvPKS genes participated in urushiol biosynthesis according to their co-expression profile and homological structure. Based on our data, we drew the skeleton of the urushiol biosynthesis pathway and provided more details at whole genome level. Moreover, the expression levels of urushiol-related genes (KCS11, KCS4, and TvPKS9, 16, 19, 22, 23) were linked to the amount of urushiol in the lacquer cultivars, further indicating those genes might participate in the urushiol biosynthesis. TvPKS genes also preferred to cluster together on the chromosome, showing that they evolved from a common ancestor gene or co-transcript to perform the same function (Figure 3B). Besides, TvPKSs could interact with enzymes in phenylpropanoid pathways, such as CHS, CHI, and 4CL (Figure 3D). This may be due to the broad promiscuity of the substrates of PKS, which enables TvPKSs to catalyze reactions on the phenylpropanoid pathway, either competitively or synergistically interacting with the 4CL, CHS, etc. This interaction might allow for fine switching of metabolic flow between different metabolic pathways. However, the speculation still requires further confirmation by enzymology or crystal structural and catalytic machinery analysis. In this study, we identified urushiol-related genes genome-wide and described generally the synthetic pathway of urushiol, updating the understanding of urushiol biosynthesis and providing candidate key enzyme genes for further genetic and enzymatic verification.

**Metabolic variations in lacquer cultivars**

The metabolomic profiling of four cultivars and the wild type of lacquer tree depicted the presence of bioactivities, such as flavonoids, phenolic acids, terpinene, lignans, and coumarins (Table S21). And we also found the chemicals related to urushiol biosynthesis, such as long-chain fatty acids, urushiols, and their analogs like pentadecatrienyl/pentadecadienyl-benzenediols (Table S22). The isolation and pharmacological identification of bioactive compounds from lacquer tree has been a popular field of research for decades, and several researchers have explored those bioactivities have antioxidant, antitumor, anti-inflammatory, antibacterial, antiviral, and neuroprotective potential (Hong et al., 2013; Kim et al., 2019a, 2019b; Choi et al., 2013; Lee et al., 2010; Cho et al., 2012; Kang et al., 2012). Fisetin, a natural flavonoid found in lacquer tree has been proved to process the excellent multiple pharmacological activities, such as antioxidant, anti-inflammatory, and anticancer effects in various cell types (Higa et al., 2003; Kim et al., 2015b; Sabarwal et al., 2017). It is reported that fisetin could induce autophagy via the crosstalk between the AMPK/mTOR and p8-dependent pathways in pancreatic cancer cells (Jia et al., 2019). And Chamcheu et al. suggested recently that fisetin has a great potential to be developed as an effective and inexpensive agent for the treatment of psoriasis and other related inflammatory skin disorders (Chamcheu et al., 2019). In addition, another flavonoid butin is potent antioxidant against oxidative stress-related diseases, such as cancer, aging, liver diseases, and diabetes (Bruselmans et al., 2005; Patil et al., 2003; Kuzu et al., 2008; Shu et al., 2009). Many other major flavonoids isolated from lacquer tree, such as sulfuretin, naringin, and pinocembrin, are known to have anti-inflammatory effects, and may have therapeutic value in preventing or delaying the
progression of rheumatoid arthritis, diabetes mellitus, and other inflammatory disease (Lee et al., 2012; Kandhare et al., 2014; Gu et al., 2017). In addition to flavonoid active components, the lacquer tree includes phenolic acids and terpenoids, such as caffeic acid and ursolic acid, with strong antioxidant, antibacterial, and antiviral properties, which are extremely healthy for us (Jayanthi and Subash, 2010; Tohmé et al., 2019).

Although we detected many active ingredients in this study, unfortunately, intermediates in the urushiol biosynthesis pathway could not be found possibly due to structural instability or low content in the sample.

The metabolic profiling is often inferred that metabolic information reflects biological endpoints more accurately than transcript or protein analysis. The cultivars have their own metabolite compositions during developmental processes or seasons. For example, the bioactive contents and the antioxidant potential of leaves of olive cultivars (Arbequina, Manzanilla, and Picual) collected in four periods of the year showed differences (Lorini et al., 2021). And the metabolism analysis of two varieties revealed that grape cultivars undergo several changes in primary metabolite concentration during berry developmental progression (Cuadros-Inostroza et al., 2016). Therefore, metabolomics is the key technology to screen the metabolic indicators of plants. Meanwhile, the metabolic markers can be used as biomarkers to identify plant species or varieties, and even the population (Sarrou et al., 2017; Sawada et al., 2019).

In our data, compared with wild lacquer, the cultivars process different contents and compositions of urushiol, and they also process their unique metabolites, such as cirsimaritin, eupatorin, persicoside, etc. Among the cultivars, we found that limocitrin-3-O-sophoroside and persicoside are the specific chemicals in HuM, and scutellarin is the unique metabolite in HoM. 7,8-Dihydroxy-5,6,4'-trimethoxyflavone and eriodictyol-7-O-(6''-malonyl) glucoside are the specific metabolites in DHP, while limocitrin-3,7-di-O-glucoside is specific to GBC. These specific metabolites could be used as metabolic markers to distinguish the four lacquer cultivars. And the special chemicals they made could be formed by long-term domestication, mutagenesis, breeding, and selection. Otherwise, the unique flavonoids in lacquer trees (such as butin, fisetin, fustin, and sulfuretin) as the pharmacological foundation could produce beneficial and safe functional diets for preventing neurological disorders (Afzal et al., 2021; Kim et al., 2015b; Zhang et al., 2011). These metabolome data provide a basis for the comprehensive utilization of the lacquer tree.

Limitations of the study

We sequenced the genome, transcriptome data, and metabolome data to provide the genetic background for the lacquer tree. We identified the candidate genes associated with the urushiol and lignin biosynthesis pathways only by analyzing the expression profile and homology blast analysis, with no experimental validation. Therefore, the function of these candidate genes should be further verified using genetic and biochemical technologies. Moreover, the metabolome data should be further developed into a stable and reliable method of identifying more lacquer cultivars and applied in the lacquer planting industry.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Genome sequencing and assembly
  - Gene prediction and annotation
  - Assembly and gene set evaluation
  - Phylogenetic analyses
  - Genome syntenic and whole-genome duplication analysis
  - Identification of polyketide synthases genes from genome sequences
  - Metabolomic profile of lacquer cultivars
  - qRT-PCR analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES
SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104512.

ACKNOWLEDGMENTS
This research was funded by National Natural Science Foundation of China (No. 31800554) and Scientific Research Project of Shaanxi Province Bureau of Traditional Chinese Medicine (grant number 610902).

AUTHOR CONTRIBUTIONS
G.B. and S.L. designed the project. Y.L., W.L, C.C., and D.L. prepared the samples and conducted experiments. T.Z., T.Z., and C.Z. performed the genome assembly and annotation and conducted bioinformatic analysis. C.C., G.B., C.Z., T.Z., and S.L. wrote and revised the manuscript.

DECLARATION OF INTERESTS
The authors declare no competing interests.

Received: December 29, 2021
Revised: April 21, 2022
Accepted: May 27, 2022
Published: July 15, 2022

REFERENCES
Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of biology. Nat. Genet. 25, 23–29. https://doi.org/10.1038/7556.
Attwood, T.K., Croning, M.D.R., Flower, D.R., Lewis, A.P., Mabey, J.E., Scordis, P., Selley, J.N., and Wright, W. (2000). PRINTS—s the database formerly known as PRINTS. Nucleic Acids Res. 28, 225–227. https://doi.org/10.1093/nar/28.1.225Aral.
Alfaz, M., Al-Abbas, F.A., Kazimi, I., Imam, S.S., Alshehri, S., Ghoneim, M.M., Almalki, W.H., Nadeem, M.S., and Sayyed, N. (2021). Fustin inhibits oxidative free radicals and inflammatory cytokines in cerebral cortex and Hippocampus and protects cognitive impairment in streptozotocin-induced diabetic rats. ACS Chem. Neurosci. 12, 4587–4597. https://doi.org/10.1021/acschemneuro.1c00712.
Auber, R.P., Suttiyut, T., McCoy, R.M., Ghaste, M., Crook, J.W., Pendleton, A.L., Widholm, J.R., and Wiseacre, J.H. (2020). Hybrid de novo genome assembly of red gromwell (Lithospermum erythrorhizon) reveals evolutionary insight into shikonin biosynthesis. Hortic. Res. 7, 82. https://doi.org/10.1038/s41438-020-0301-9.
Bai, G.Q., Jia, Y., Li, W.M., Chen, H., Li, B., and Li, S.F. (2018). De novo assembly and analysis of sumac (Toxicodendron vernicifluum (Stokes) F.A. Barkley) transcriptomes provides insights into the biosynthesis of urushiol. Mol. Biol. 52, 660–667. https://doi.org/10.1007/s10026-018-00505-9.
Bao, W., Kojima, K.K., and Kohany, O. (2015). RefSeqUpdate, a database of repetitive elements in eukaryotic genomes. Mobile DNA 6, 11. https://doi.org/10.1186/s13100-015-0041-9.
Barakat, A., Yassin, N.B.M., Park, J.S., Choi, A., Herr, J., and Carlson, J.E. (2011). Comparative and phylogenomic analyses of cinnamoyl-CoA reductase and cinnamyl-CoA-reductase-like gene family in land plants. Plant Sci. 181, 249–257. https://doi.org/10.1016/j.plantsci.2011.05.012.
Bateman, A., Cohn, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., and Sonnhammer, E.L.L. (2004). The Pfam protein families database. Nucleic Acids Res. 32, D138–D141. https://doi.org/10.1093/nar/gkh121.
Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27, 573–580. https://doi.org/10.1093/nar/27.2.573.
Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and genomewise. Genome Res. 14, 988–995. https://doi.org/10.1101/gr.1865504.
Boeckmann, B., Barocho, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O’Donovan, C., and Phan, I. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. 31, 365–370. https://doi.org/10.1093/nar/gkg095.
Bonnehoh, K., Poch, G., and Metz, M. (2018). Severe contact dermatitis caused by urushiol in Japanese lacquer. Contact Dermatitis 80, 55–56. https://doi.org/10.1111/cod.13104.
Brusselmans, K., Vrolix, R., Verhoeven, G., and Swinnen, J.V. (2005). Induction of cancer cell apoptosis by flavonoids is associated with their ability to inhibit fatty acid synthase activity. J. Biol. Chem. 280, 5636–5645. https://doi.org/10.1074/jbc.M401772200.
Cantarel, B.L., Korf, I., Robb, S.M., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez-Alvarado, A., and Yandell, M. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18, 188–196. https://doi.org/10.1101/gr.43907.
Chamcheu, J.C., Esnault, S., Adhami, V.M., Noll, A.L., Banang-Mbeumi, S., Roy, T., Singh, S.S., Huang, S., Koussoulas, K.G., and Mukhtar, H. (2019). Fisetin, a 3, 7, 3’-4’-tetrahydroxyflavone inhibits the PI3K/Akt/mTOR and MAPK pathways and ameliorates psoriasis pathology in 2D and 3D organotypic human inflammatory skin models. Cells 8, 1089. https://doi.org/10.3390/cells8091089.
Chao, N., Yu, T., Hou, C., Liu, L., and Zhang, L. (2021). Genome-wide analysis of the lignin toolbox for morus and the roles of lignin related genes in response to zinc stress. PeerJ 9, e11964. https://doi.org/10.7717/peerj.11964.
Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., Li, Y., Ye, J., Yu, C., Li, Z., et al. (2018). SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. Gigascience 7, 1–6. https://doi.org/10.1093/gigascience/gqx120.
Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., Yu, S., Xiong, L., and Luo, J. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. Mol. Plant. 6, 1769–1780. https://doi.org/10.1093/mp/sss080.
Cheng, X., Li, M., Li, D., Zhang, J., Jin, Q., Sheng, L., Cai, Y., and Lin, Y. (2017). Characterization and analysis of CCR and CAD gene families at the whole-genome level for lignin synthesis of stone cells in pear (Pyrus brechtneri) fruit. Biol. Open 6, 1602–1613. https://doi.org/10.1242/bio.026997.
Cho, J.Y., Park, K.Y., Kim, S.J., Oh, S., and Moon, J.H. (2015). Antimicrobial activity of the synthesized non-allergenic urushiol derivatives. Biosci. Biotechnol. Biochem. 79, 1915–1918. https://doi.org/10.1007/s10529-015-10614-8.
Cho, N., Choi, J.H., Yang, H., Jeong, E.J., Lee, K.Y., Kim, Y.C., and Sung, S.H. (2012). Neuroprotective and anti-inflammatory effects of
flavonoids isolated from Rhus verniciflua in neuronal H722 and microglial BV2 cell lines. Food Chem. Toxicol. 50, 1940–1945. https://doi.org/10.1016/j.fct.2012.03.052.

Choi, H.S., Seo, H.S., Kim, S.R., Choi, Y.K., Jang, B.H., Shin, Y.C., and Ko, S.G. (2014). Anti-inflammatory and anti-proliferative effects of Rhus verniciflua Stokes in RAW264.7 cells. Mol. Med. Rep. 9, 311–315. https://doi.org/10.3892/mmr.2013.1775.

Boachon, B., Buell, C.R., Crisovan, E., Dudareva, N., Garcia, N., Godden, G., Henry, L., Kamileen, M.O., Kates, H.R., Kilgore, M.B., et al. (2018). Phylogenomic mining of the mints reveals multiple mechanisms contributing to the evolution of chemical diversity in Lamiaceae. Mol. Plant 11, 1084–1096. https://doi.org/10.1016/j.molp.2018.06.002.

Cuadros-Inostroza, A., Ruiz-Lara, S., González, E., Eckhardt, A., Willmitzer, L., and Peña-Cortés, H. (2016). GC-MS metabolic profiling of Cabernet Sauvignon and Merlot cultivars during grapevine bunch development and network analysis reveals a stage- and cultivar-dependent connectivity of primary metabolites. Metabolomics 12, 39. https://doi.org/10.1007/s11306-015-0927-z.

Cui, J., Lu, Z., Wang, T., Chen, G., Mostafa, S., Ren, H., Liu, S., Fu, C., Wang, L., Zhu, Y., et al. (2021). The genome of Medicago polymorpha provides insights into its edibility and nutritional value as a vegetable and forage legume. Hortic. Res. 8. 47. https://doi.org/10.1038/s41438-021-00483-5.

De Bie, T., Cristiani, N., Demuth, J.P., and Hahn, M.W. (2006). CAFE: a computational tool for the study of gene family evolution. Bioinformatics 22, 1249–1251. https://doi.org/10.1093/bioinformatics/bti907.

De Coster, W., D’Hert, S., Schultz, D.T., Cruts, M., and Van Breukelen, C. (2018). NanoPack: visualizing and processing long read sequencing data. Bioinformatics 34, 2666–2669. https://doi.org/10.1093/bioinformatics/bty149.

Draper, W.M., Wijekoon, D., McKinney, M., Behniwal, P., Perera, S.K., and Flessell, C.P. (2002). Atmospheric pressure ionization LC-MS-MS determination of urushiol congeners. J. Agric. Food Chem. 50, 1852–1858. https://doi.org/10.1021/jf011364t.

Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Høegger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., and Aiden, E.L. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science 356, 92–95. https://doi.org/10.1126/science.aai3327.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M., Lander, E.S., and Aiden, E.L. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. 3, 95–98. https://doi.org/10.1016/j.cels.2016.07.002.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1797–1800. https://doi.org/10.1093/nar/gkh340.

Edgar, R.C., and Myers, E.W. (2005). PILER identification and classification of genomic repeats. Bioinformatics 21, i152–i158. https://doi.org/10.1093/bioinformatics/bti1003.

ElSoHly, M.A., Adawadkar, P.D., Ma, C.Y., and Tumer, C.E. (1982). Separation and characterization of poison ivy and poison oak urushiol components. J. Nat. Prod. 45, 532–538. https://doi.org/10.1021/np50023a004.

Emms, D.M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 16, 137. https://doi.org/10.1186/s13059-015-0721-2.

Fan, G., Liu, X., Sun, S., Shi, C., Du, X., Han, K., Yang, B., Fu, Y., Liu, M., Seim, I., et al. (2020). The chromosome level genome and genome-wide association study for the agronomic traits of Panax Notoginseng. (Science 23, 101538. https://doi.org/10.1126/sciadv.abc01.

Cao, X., Li, S., Wang, C., and Lu, M. (2007). Distributed regularity and anatomical feature of laticifer canals of Toxicodendron vernicifloum (Stokes) F. A. Barkley. Acta Bot. Boreal.-Occid. Sin. 23, 885–890. https://doi.org/10.11872/2075160055-7.

Goujon, T., Sibout, R., Eudes, A., Mackay, J., and Jouann, L. (2003). Genes involved in the biosynthesis of lignin precursors in Arabidopsis thaliana. Plant Physiol. Biochem. 41, 677–687. https://doi.org/10.1016/j.gebre.2013.10.095.

Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R., and Bateman, A. (2004). Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res. 33, D121–D124. https://doi.org/10.1093/nar/gkj481.

Gu, X., Zhang, Q., Gu, D., Shen, H., and Zhu, Z. (2017). Pincombinerin alleviates allergic airway inflammation via inhibition of NF-kB pathway in mice. Int. Immunopharm. 33, 90–95. https://doi.org/10.1016/j.intimp.2017.10.005.

Guo, W., Wu, H., Zhang, Z., Yang, C., Hu, L., Shi, X., Jian, S., Shi, H., and Huang, Y. (2017). Comparative analysis of transcriptomes in rhizophoraceae provides insights into the origin and adaptive evolution of mangrove plants in intertidal environments. Front. Plant Sci. 8, 795. https://doi.org/10.3389/fpls.2017.00795.

Harigaya, S., Honda, T., Rong, L., and Miyakoshi, T. (2007). Enzymatic dehydrogenative polymerization of urushiol in fresh exudates from the lacquer tree, Rhus verniciflua DC. J. Agr. Food Chem. 55, 2201–2208. https://doi.org/10.1021/jf063161g.

Higa, S., Hirano, T., Kotani, M., Matsumoto, M., Fujita, A., Suemura, M., Kawase, I., and Tanaka, T. (2003). Fisetin, a flavonol, inhibits th2-type cytokine production by activated human basophils. J. Allergy Clin. Immunol. 111, 1299–1306. https://doi.org/10.1016/j.jaci.2003.10.046.

Honda, T., Lu, R., Sakai, R., Ishimura, T., and Miyakoshi, T. (2008). Characterization and comparison of Asian lacquer saps. Prog. Org. Coat. 61, 68–75. https://doi.org/10.1016/j.porgcoat.2007.09.003.

Hong, S.H., Suk, K.T., Choi, S.H., Lee, J.W., Sung, H.T., Kim, C.H., Kim, E.J., Kim, M.J., Han, S.H., Kim, M.Y., et al. (2013). Anti-oxidant and natural killer cell activity of Korean red ginseng (Panax ginseng) and urushiol (Rhus verniciflua Stokes) on non-alcoholic fatty liver disease of rat. Food Chem. Toxicol. 55, 586–591. https://doi.org/10.1016/j.fct.2013.01.022.

Jayanthi, R., and Subash, P. (2010). Antioxidant effect of caffic acid on oxysteroylcyline induced lipid peroxidation in albinor rats. Indian J. Clin. Biochem. 25, 371–375. https://doi.org/10.1007/s11291-010-0052-8.

Jia, S., Xu, Z., Zhou, S., Shen, Y., Ding, G.P., Cao, L., and Gao, L.P. (2019). Fisetin induces a reverse transcriptase inhibitory activity from the bark of Rhus verniciflua Stokes against fish pathogenic viruses in Vitro. J. Microbiol. 50, 293–300. https://doi.org/10.1007/s12275-012-2068-9.
simple phenolic acids in plants. Phytochem. Rev. 19, 865–806. https://doi.org/10.1007/s11101-020-009689-2.

Aziz, M., Sturtevant, D., Winston, J., Collakova, E., Jelesko, J., and Chapman, K. (2017). MALDI-MS imaging of urushiol in poison ivy stem. Molecules 22, 711. https://doi.org/10.3390/molecules22050711.

Nie, Z.L., Sun, H., Meng, Y., and Wen, J. (2009). Phylogenetic analysis of Toxicodendron (Anacardiaceae) and its biogeographic implications on the evolution of north temperate and tropical intercontinental disjunctions. J. Syst. Evol. 47, 416–430. https://doi.org/10.1111/j.1759-6831.2009.00045.x.

Patil, C.S., Singh, V.P., Satyanarayan, P.S., Jain, N.K., Singh, A., and Kulkarni, S.K. (2003). Protective effect of flavonoids against aging- and lipopolysaccharide-induced cognitive impairment in mice. Pharmacology 69, 59–67. https://doi.org/10.1155/000072357.

Patel, S., Robben, M., Fennell, A., Londo, J.P., Alahakoon, D., Villegas-Diaz, R., and Swaminathan, P. (2009). Draft genome of the native American cold hardy grapevine Vitis riparia Michx. ‘Manitoba 37’. Hortic. Res. 13, 71–83. https://doi.org/10.1038/s41438-020-0316-2.

Chang, T.-C., Mendell, J.T., and Salzberg, S.L. (2002). Downregulation of cinnamoyl CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. Plant J. 33, 290–295. https://doi.org/10.1046/j.1365-313X.1998.00014.x.

Price, A.L., Jones, N.C., and Pevzner, P.A. (2005). De novo identification of repeat families in large genomes. Bioinformatics 21, 3351–3358. https://doi.org/10.1093/bioinformatics/bti108.

Rojas-Munoz, E., Lepoittevin, J.P., Pujol, R.M., and Gimenez-Arnau, A. (2012). Allergic contact dermatitis to plants: understanding the chemistry will help our diagnostic approach. Actas Dermo Sifiliogr. 103, 456–477. https://doi.org/10.1016/j.ad.2012.07.006.

Sabanwal, A., Agarwal, R., and Singh, R.P. (2017). Fisetin inhibits cellular proliferation and induces mitochondria-dependent apoptosis in human gastric cancer cells. Mol. Carcinog. 56, 499–514. https://doi.org/10.1002/mc.22252.

Sarrou, E., Ganopoulos, I., Xanthopoulos, A., Masoura, D., Mavromatis, S., Madesis, P., Pavromatis, A., and Chatzopoulou, P. (2017). Genetic diversity and metabolic profile of Salvia officinalis populations: implications for advanced breeding strategies. Planta 246, 201–215. https://doi.org/10.1007/s00425-017-2666-z.

Sawada, Y., Sato, M., Okamoto, M., Masuda, J., Yamaki, S., Takahashi, I., Tanokashira, Y., Kаяmoto, S., Ohira, A., Abe, T., and Hirai, M.Y. (2019). Metabolome-based discrimination of chrysanthemum cultivars for the efficient generation of flower color variations in mutation breeding. Metabolomics 15, 118. https://doi.org/10.1007/s11306-019-1576-7.

Servant, F., Bru, C., Carrère, S., Courcelle, E., Gouzy, J., Peyruc, D., and Kahn, D. (2003). ProDom: automated clustering of homologous domains. Brief. Bioinform. 3, 246–251. https://doi.org/10.1093/bib/3.3.246.

Shen, H., Mazzari, M., Hisano, H., Escamilla-Trevino, L., Fu, C., Pu, Y., Ruds, M.R., Tang, Y., Xiao, J., Jackson, L., et al. (2013). A genomics approach to deciphering lignin biosynthesis in switchgrass. Plant Cell 25, 4342–4361. https://doi.org/10.1105/tpc.113.118828.

Xiao, X., An, Z., Zhou, B., Zhang, B., Tan, X., et al. (2016). PANTHER: a library of protein families and subfamilies indexed by function. Genome Res. 13, 2129–2141. https://doi.org/10.1101/gr.772403.

Tohme, M., Giménez, M., Peralta, A., Colombo, M.I., and Delguit, L.R. (2019). Ursolic acid: a novel antiviral compound inhibiting rotavirus infection in vitro. Int. J. Antimicrob. Agents 54, 601–609. https://doi.org/10.1016/j.ijantimicag.2019.07.015.

Vaser, R., Sovic, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 27, 737–746. https://doi.org/10.1101/gr.214270.116.

Vurtle, G.W., Sedlacek, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J., and Schatz, M.C. (2017). GenomeScope: fast, reference-free genome profiling from short reads. Bioinformatics 33, 2202–2204. https://doi.org/10.1093/bioinformatics/btx153.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelbel, A., Sakhikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9, e112963. https://doi.org/10.1371/journal.pone.0112963.

Wang, L., He, N., Li, Y., Fang, Y., and Zhang, F. (2020a). Complete chloroplast genome sequence of Chinese lacquer tree (Toxicodendron vernicifluum, Anacardiaceae) and its phylogenetic significance. BioMed Res. Int. 2020, 9014873. https://doi.org/10.1155/2020/9014873.

Wang, P., Luo, Y., Huang, J., Gao, S., Zhu, G., Zhang, G., Gai, J., Yang, M., Zhu, M., Zhang, H., et al. (2020b). The genome evolution and domestication of tropical fruit mango. Genome Biol. 21, 60. https://doi.org/10.1186/s13059-020-01959-8.

Wang, Y., Tang, H., DeBerry, J.D., Tan, X., Li, J., Wang, X., Lee, T.-F., Jin, H., Marler, B., Guo, H., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 40, e49. https://doi.org/10.1093/nar/gkr1293.

Wang, Y., Xin, H., Fan, P., Zhang, J., Liu, Y., Dong, Y., Wang, Z., Yang, Z., Zhang, Q., Ming, R., et al. (2021). The genome of Shanputao (Vitis amurensis) provides a new insight into cold tolerance of grapevine. Plant J. 105, 1495–1506. https://doi.org/10.1111/tpj.15127.

Watanabe, H., Fujimoto, A., Nishida, J., Ohishi, T., and Takahara, A. (2016). Biobased polymer coating using catechol derivative urushiol. Langmuir 32, 4619–4623. https://doi.org/10.1021/acs.langmuir.6b00484.

Weisberg, A.J. (2014). Investigations into the Molecular Evolution of Plant Terpene, Alkaloid, and Unavoid Biosynthetic Enzymes Virginia Polytechnic Institute and State University).

Wolfe, K.H. (2001). Yesterday's polyploids and the mystery of diploidization. Nat. Rev. Genet. 21, 333–341. https://doi.org/10.1038/sj.embor.7400709.

Wu, T.D., and Watanabe, C.K. (2005). GMAP: a genomic mapping and alignment program for
## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Biological samples** | | |
| The total DNA and RNA of wild T. vernicifluum and the varieties | This study | NA |
| **Critical commercial assays** | | |
| TRizol™ | Invitrogen | 10296010 |
| DNA isolation | CTAB | - |
| NEBNext Ultra II End RepairdA-Tailing Module | NEB | E7546L |
| **Deposited data** | | |
| Assembly genome | This paper | GenBank: PRJNA836312 |
| **Software and algorithms** | | |
| SOAPnuke v.1.5.6 | Chen et al., 2018 | https://github.com/BGI-flexlab/SOAPnuke |
| NanoFilt | De Coster et al., 2018 | https://github.com/wdecoster/nanofilt |
| GenomeScope v1.0 | Vurture et al., 2017 | https://github.com/schatzlab/genomescope |
| Jellyfish v2.0 | Marcias and Kingsford, 2012 | https://github.com/gmarciais/Jellyfish |
| SMARTdenovo | Liu et al., 2021 | https://github.com/ruanjue/smartdenovo |
| Racon v1.3.3 | Vaser et al., 2017 | https://github.com/sovic/racon |
| Pilon v.1.225 | Walker et al., 2014 | https://github.com/broadinstitute/pilon |
| Juicer v.1.5 | Durand et al., 2016 | https://github.com/aidenlab/juicer |
| 3D-DNA, v.180922 | Dudchenko et al., 2017 | https://github.com/aidenlab/3d-dna |
| GMAP | Wu and Watanabe, 2005 | https://github.com/julianehering/GMAP-GSNAP |
| HIsAT v2.1.1 | Kim et al., 2015a | http://daehwankimlab.github.io/hiSAT2/ |
| StringTie v1.3.3b | Pertea et al., 2015 | https://github.com/gpertea/stringtie |
| RSEM v1.2.12 | Li and Dewey, 2011 | https://github.com/deweylab/RSEM |
| DEseq2 | Love et al., 2014 | https://bioconductor.org/packages/release/bioc/html/DESeq2.html |
| RepeatMasker v.4.0.7 and RepeatProteinMask v.4.0.7 | Tarailo-Graovac and Chen, 2009 | http://www.repeattmasker.org/cgi-bin/RepeattProteinMaskRequest |
| Repbase database v21.12 | Bao et al., 2015 | http://www.girinst.org/repbase |
| LTR-FINDER | Xu and Wang, 2007 | http://tife.fudan.edu.cn/trr_finder/ |
| PilER | Edgar and Myers, 2005 | http://www.drive5.com/piler |
| RepeatScout | Price et al., 2005 | http://bx.ucsd.edu/repeatscout/ |
| TRF v.4.09 | Benson, 1999 | https://tandem.bu.edu/trf/trf.html |
| Augustus | Stanke et al., 2006 | https://github.com/Gaius-Augustus/Augustus |
| SNAP | Korf, 2004 | https://github.com/KorfLab/SNAP |
| MAKER pipeline | Cantarel et al., 2008 | https://www.yandell-lab.org/software/maker.html |
| GeMoMa v1.6 | Keilwagen et al., 2019 | http://www.jstacs.de/index.php/GeMoMa |
| tRNAscan-SE v1.3.1 | Lowe and Eddy, 1997 | http://lowelab.ucsc.edu/trfANACan-SE |
| Rfam v12.0 database | Griffiths-Jones et al., 2005 | https://rfam.xfam.org/ |
| BUSCO v4.0.6 | Simão et al., 2015 | https://busco.ezlab.org/ |
| BLASTp | Kent, 2002 | https://blast.ncbi.nlm.nih.gov/Blast.cgi |

(Continued on next page)
**RESOURCE AVAILABILITY**

**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Sifeng Li (lisf60@sina.com).

**Materials availability**

This study did not generate new unique reagents.

**Data and code availability**

- The accession number for the genome assembly and raw reads reported in this paper is GenBank: PRJNA836312.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the Lead contact upon request.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

This study does not include experiments model or subjects.

**METHOD DETAILS**

**Genome sequencing and assembly**

*Plant materials and genome sequencing*

We extracted high-quality genomic DNA from fresh leaves of a diploid female *T. vernicifluum* ‘Gaobachi’ grown in Xi’an Botanical Garden, Xi’an, Shanxi Province (E109°14′3.65″, N34°12′31.32″). The libraries with short insert sizes of 350 bp and long insert sizes of about 20 kb were constructed and sequenced according to the standard protocols on the MGISEQ-2000 platform and Oxford Nanopore platform, respectively. DNBSEQ clean data was filtered by SOAPnuke v.1.5.6 (Chen et al., 2018), and we were trimmed and filtered ONT reads by NanoFilt (De Coster et al., 2018), with base quality score < 7 and reads which shorter than 5 kb.

For building a Hi-C library, genomic DNAs were fixed with formaldehyde and were sheared by a restriction enzyme (MboI). The library was then sequenced by MGISEQ-2000 platform and filtering by SOAPnuke v.1.5.6. A total of 488,159,058 read pairs with 73.22 Gb of 150PE Hi-C data were generated to acquire a chromosome-level assembly of the genome.
De novo assembly
Genome size, repeat fraction and heterozygosity of *T. vernicifluum* were estimated by using GenomeScope v1.0 (Vurture et al., 2017) based on the k-mer histograms provided by Jellyfish v2.0 (Marcais and Kingsford, 2012) with 17-mer. After corrected and trimmed by Canu v1.7, the ONT long reads were de novo assembled by using SMARTdenovo (https://github.com/ruanjue/smartdenovo) with the parameters `-k 26 -z 10 -U -1 -m 0.6 -A 1000`. Next, three rounds of minimap2 (Li, 2018)-Racon v1.3.3 (Vaser et al., 2017) and one round of MEDAKA (https://github.com/nanoporetech/medaka) were applied to polish the assembly with uncorrected ONT sequence data, and finally Pilon v.1.225 (Walker et al., 2014) was employed to calibrate the consensus contigs using DNBSEQ short reads.

Then, Hi-C reads were mapped to draft genome assembly and processed to generate interaction matrices by Juicer v.1.5 (Durand et al., 2016). 3D de novo assembly (3D-DNA, v.180922) pipeline (Dudchenko et al., 2017) was performed to anchor primary contigs into chromosome-level assembly by using the output of Juicer pipeline.

Transcriptome sequencing and gene expression analysis
RNAs from leaves, roots, flowers and drupes from *T. vernicifluum* ‘Gaobachi’ were extracted for full-length transcriptome sequencing using the PacBio Sequel II platform. PacBio ISO-Seq3 pipeline (https://github.com/PacificBiosciences/IsoSeq) was used to obtain full-length non-chimeric (FLNC) transcripts via ccs, classify, cluster and polish stages. The high-quality transcripts were mapped against the assembled genome using GMAP (Wu and Watanabe, 2005), and collapsed to reduce redundancy using CupcakeToFU (https://github.com/Magdoll/cDNA_Cupcake/wiki/Cupcake-ToFU) scripts.

Five varieties of *T. vernicifluum* were sampled for RNA-seq analysis, include ‘Gaobachi’ (GBC), ‘Dahongpao’ (DHP), ‘Huangmaoguizhou’ (HuM), ‘Hongmaoguizhou’ (HoM) and wild variety, each has three biological repeats, libraries of insert size 300–400 bp were prepared according to the manufacturer’s protocol of DNBSEQ sequencing platform and paired-end sequenced by MGISEQ-2000 platform. The reads filtering by SOAPnuke were mapped to assembled genome using HISAT v2.1.2 (Kim et al., 2015a) and assembled by StringTie v1.3.3b (Pertea et al., 2015). Gene expression levels in terms of fragments per kilobase per million reads mapped (FPKM) values were calculated by RSEM v1.2.12 (Li and Dewey, 2011) and differentially expressed genes (DEGs) were detected by DEseq2 (Love et al., 2014).

Gene prediction and annotation
Repetitive elements in the *T. vernicifluum* genome were identified through a combination of homology-based and de novo approaches. RepeatMasker v.4.0.7 and RepeatProteinMask v.4.0.7 (Tarailo-Graovac and Chen, 2009) were used to detect repeats by aligning against Repbase database v21.12 (Bao et al., 2015). LTR-FINDER (Xu and Wang, 2007), Piler (Edgar and Myers, 2005) and RepeatScout (Price et al., 2005) were employed to build transposable element (TE) consensus sequences as a de novo TE library, and TRF v.4.09 (Benson, 1999) was used to obtain tandem repetitive sequences. RepeatMasker v.4.0.7 was then used to discover and identify repetitive sequences with the combined library of the de novo TE.

MAKER pipeline (Cantarel et al., 2008) was employed twice to annotate protein-coding genes in genome with comprehensive evidence from ab initio prediction, homology-based and transcriptome-based prediction. Protein sequences of *Acer yangbiense* (GigaDB), *Anacardium occidentale* (Phytozome v13), *Arabidopsis thaliana* (TAIR 10), *Citrus sinensis* (Phytozome v13), *Dimocarpus longan* (GigaDB) and *Pistacia vera* were used as homology-based evidence. Non-redundant isoforms were used as transcriptome-based evidence. At first round, the genome was annotated to gain genes by Maker pipeline with proteins and transcripts evidence. The above result was used to train the parameters for Augustus (Stanke et al., 2006) and SNAP (Korf, 2004). The second round, we combined the models trained by Augustus and SNAP with protein sequences of five species and isoforms to predict the gene set. Homologous genes from *Pistacia vera* was predicted by GeMoMa v1.6 (Keilwagen et al., 2019) and compared with the gene set generated by MAKER. Then the gene models not predicted by MAKER were integrated into the final gene set.

For the annotation of non-coding RNA, we detected tRNA by trRNAscan-SE v1.3.1 (Lowe and Eddy, 1997), and identified other noncoding RNA, including miRNA, rRNA and snRNA, by searching Rfam v12.0 database (Griffiths-Jones et al., 2005).
Gene functional annotation was performed based on consensus of sequence and domain. The protein sequences were aligned to NCBI Non-Redundant Protein Sequence (NR) databases, Kyoto Encyclopedia of Genes and Genomes (KEGG v89.0; (Kanehisa et al., 2017)), SwissProt and TrEMBL (Uniprot release 2020-06) (Boeckmann et al., 2003) with BLASTp. The domains were searched and predicted by using InterProScan v5.11-55.0 (Zdobnov and Apweiler, 2001) with publicly available databases including PANTHER (Thomas et al., 2003), Pfam (Bateman et al., 2004), PRINTS (Attwood et al., 2000), ProDom (Servant et al., 2002), PROSITE profiles (Sigrist et al., 2010), and SMART (Letunic et al., 2012). Gene ontology (GO) terms (Ashburner et al., 2000) for each gene were predicted from the InterPro descriptions.

Assembly and gene set evaluation
DNBSEQ clean reads and ONT corrected reads were mapped against the *T. vernicifluum* genome with BWA-MEM (Li, 2013) and minimap2 (Li, 2018) to evaluate the quality of our assembly. BUSCO v4.0.6 (Sima˜o et al., 2015) with parameter `-l embryophyta_odb10` was applied to evaluate completeness of the assembly and gene set.

Phylogenetic analyses
We collected *T. vernicifluum* and other fourteen species, including the species used for gene annotation as well as Carica papaya, Oryza sativa, Prunus persica, Solanum lycopersicum, Theobroma cacao, Vitis vinifera downloaded from Phytozome v13, Sclerocarya birrea (GigaDB) and Mangifera indica (Wang et al., 2020a, 2020b) to identify gene families. After removing alternative splicing sequences, we performed Orthofinder (Emms and Kelly, 2015) to obtain orthologous and paralogous genes of these species. Then RaxML (Stamatakis, 2006) was applied to construct the phylogenetic tree by using the single copy orthologous gene families with the GTRGAMMA model. MCMCtree program in PAML v4.9j (Yang, 2007) was performed to estimate the divergence time between fifteen species in the phylogenetic tree with the REV substitution model. Four calibration time points based on the TimeTree database (http://www.timetree.org) were used as references, including *A. thaliana* and *A. yangbiense* (96- 104 Mya), *A. yangbiense* and *D. longan* (53- 104 Mya), *A. yangbiense* and *C. sinensis* (81- 103 Mya), and *A. occidentale* and *C. sinensis* (62- 84 Mya).

CAFE (De Bie et al., 2006) was employed to investigate gene family expansion and contraction under a maximum likelihood framework, single copy orthologous gene families and estimated divergence time between different species were used as input files. Families with p-value < 0.05 were regarded as significant expansion or contraction and were used for enrichment analysis.

Genome synteny and whole-genome duplication analysis
We applied MCScanX v1.5.2 (Wang et al., 2012) to detect synteny among *T. vernicifluum* and other species in Anacardiaceae, and calculated the 4DTv of syntenic blocks. The WGDI (Sun et al., 2021) pipeline also used to estimate synonymous substitution per site (Ks) of these species. Then, the WGD events were evaluated with the 4DTv and Ks distribution.

Identification of polyketide synthases genes from genome sequences
Reference sequences of Polyketide synthases (PKS) genes were download from NCBI. Sequences of PKS genes were aligned from *A. occidentale*, *A. thaliana*, *M. indica*, *P. vera*, *O. sativa*, *S. birrea* and *T. vernicifluum* by using TBLASTn with an E-value of 1 × 10^{-5}, and the blasted hits were conjoined by SOLAR v0.9 (Yu et al., 2006). We predicted the structure of PKS genes by GeneWise v2.4 (Birney et al., 2004), and extended the 3' and 5' direction in potential PKS genes that incomplete. PKS genes with interrupting stop codons or frameshifts were classified as pseudogenes, and the genes with a partially intact were classified as partial genes. The PKS domain numbered PF00195 and PF02797 obtained from the Pfam database was used as a standard sequence, and then all candidate sequences were analyzed in the Pfam database by InterProScan. The potential genes from start codon to stop codon and contain N-terminal domain and C-terminal domain were considered as intact PKS genes. We converted coding sequences (CDS) to protein sequences and used Muscle v3.8.31 (Edgar, 2004) to perform multiple sequence alignments. RaxML was employed to construct gene family phylogenetic tree with the PROTGAMMAAUTO model.
Metabolomic profile of lacquer cultivars
We collected the phloem of the wild lacquer (TV) and cultivars HoM, DHP, GBC, and HuM for metabolomic analysis. Weighed 100-mg samples and added 1.2 mL of 70% methyl alcohol extraction solvent to each sample, and vortexed every 30 min for 30 s, and stored the extractions at 4°C overnight. Then, the samples were centrifuged for 10 min at 12,000 rpm and filtrated by 0.22 μM filter membrane. Next, 4 μL of supernatant was injected into an Agilent SB-C18 column (1.8 μm, 2.1 × 100 mm) for ultraperformance liquid chromatography-electrospray ionization-mass spectrometry (UPLC-MS/MS) analysis. For qualitative analysis, we used the MVDB V2.0 database of Wuhan Maiteville Biotechnology Co., Ltd. (Wuhan, China). Based on the stepwise MIM–EPI (multiple ion monitoring-enhanced product ions) to build the commercially available standard Metabolites Database (Metware Biotechnology Co., Ltd. Wuhan, China) and the public metabolite databases such as MassBank, KNAPSacK, HMDB, MoToDB, and METLIN, qualitative analysis of the metabolite data was performed.

qRT-PCR analysis
Nine genes were involved in urushiol and lignin pathways randomly chosen to validate the RNA-seq data. Total RNA samples were same as the RNA-seq analysis. The First-strand cDNA was reverse transcribed using the PrimeScript™ RT reagent Kit with gDNA Eraser (Takara, Dalian, China). qRT-PCR was performed using SYBR Premix Ex Taq™ (Takara, Dalian, China). All samples were normalized to TvActin to determine the isoform expression level. qRT-PCR primers were designed with PRIMER PREMIER 6 (PREMIER Biosoft, http://www.premierbiosoft.com/primerdesign/), and their specificity was verified by PCR (Table S24). The conditions for the qRT-PCR amplification were as follows: incubation at 95°C for 30 sec, 35 cycles of 95°C for 5 sec and 60°C for 30 sec. The specificity of the primer amplicons was tested using a melting curve analysis. The amplification was carried out in triplicate with LightCycler®96 instrument (Roche Diagnostics, Switzerland) and the expression levels of the candidate genes were calculated using the 2^–ΔΔCt method.

QUANTIFICATION AND STATISTICAL ANALYSIS
The quantitative analysis of metabolites used multiple reaction monitoring (Chen et al., 2013). Significantly different metabolites between groups were determined by variable importance in projection (VIP) ≥ 1 and an absolute Log2FC (fold change) ≥ 1 (Jiang et al., 2020). Linear discriminate analysis (LDA) effect size analysis of top 10 differential metabolites among groups and LDA scores and significance of p ≤ 0.05 as determined by Wilcoxon’s signed-rank test.

ADDITIONAL RESOURCES
This study does not include additional resources.