Transcriptome-based Identification and Molecular Evolution of Cytochrome P450 Genes and Expression Profiling during sycamore lace bug Corythucha ciliata Feeding on London planetree Platanus acerifolia

Abudurusuli Tusun
Xinjiang Normal University
Fengqi Li
Guizhou University
Youssef Dewer
Agricultural Research Center
chunyan wu (kinghour@163.com)
Guizhou University

Research Article

Keywords: London planetree, P450, gene family, gene expression, evolutionary analysis

Posted Date: April 26th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1531493/v1

Version of Record: A version of this preprint was published at Brazilian Journal of Botany on November 30th, 2023. See the published version at https://doi.org/10.1007/s40415-023-00959-9.

Abstract

Background Plant cytochrome P450s are important in the biosynthesis of essential physiological compounds. However, the gene family of P450s in london planetree have not yet been reported.

Objective To explore the sequence characteristics and expression pattern of P450s in london planetree.

Methods The phylogenetic studies, evolutionary analysis,gene expression profiling, Weighted correlation network analysis of P450 genes in london planetree were conducted. The jasmonic acid pathway key synthetase gene AOS was analyzed by qRT-PCR, cloned, homology modeled, and molecular docking analysis.

Results The london planetree P450 genes were divided into 8 clans, 38 families, and 50 subfamilies. Among them, 48 genes showed a strong negative selec-tion of P450 genes between london planetree and Arabidopsis thaliana. Nine P450 genes were sig-nificantly up-regulated at 24 h and 48 h after feeding induction by sycamore lace bug *Corythucha ciliata*. After Weighted correlation network analysis (WGCNA), the P450 genes were identified by a co-expression network mod-ule after feeding, and we identified CYP74A187 (T107931_c2_g1), CYP94A140 (T107552_c0_g3), and CYP79A208 (T94418_c0_g1) as the three hub genes for the expression module. An additional nine key P450 genes were verified by qPCR. The jasmonic acid (JA) pathway key synthetase gene AOS (CYP74A187) was cloned, homology modeled, and subjected to molecular docking analysis.

Conclusion This study identified the P450 gene family members of london planetree, and its expression, evolution, and anti-insect defense response. These data increased our understanding of the bio-logical stress of london planetree.

Introduction

The london planetree *Plantanus acerifolia* has strong adaptability, with noise reduction, anti-pollution and bactericidal ability (Zhang et al. 2011). It has become the first choice of common urban and plant greening tree species (Fengqi Li et al. 2019). The sycamore lace bug *Corythucha ciliata* (Say) is an invasive insect species and a pest of *P. acerifolia*, *Platanus occidentalis*, and *Plantanus orientalis*. Sycamore lace bug mainly damages leaves by sucking the sap, causing the leaves to fall off, and weakens or kills the tree (Li et al. 2017; Feng-Qi et al. 2018; Halbert and Meeker 1983). The interaction mechanism of london planetree defense is poorly known and little research exists on the insect resistant mechanism of london planetree.

Plant P450s proteins are involved in several biochemical pathways. Many primary and secondary metabolites such as phenylpropanoids, terpenoids, alkaloids, lipids, glucosinolates, and cyanogenic glycosides, as well as plant hormones, are produced from these biochemical synthesis pathways (Mizutani and Ohta 2010). More than 28,500 P450 sequences have been stored in The Cytochrome P450 Homepage (http://drnelson.uthsc.edu/CytochromeP450.html)(Nelson 2011). Cytochrome P450 is the

third-largest gene family in *Arabidopsis thaliana* with 245 known genes (Nelson and Werck-Reichhart 2011).

Cytochrome P450s are closely related to insect resistance. The P450 gene superfamily involves the biosynthesis of many anti-insect hormones and toxin compounds in plants and includes a variety of insect-resistant chemicals such as jasmonic acid and benzoxazinoids (Schuler 2011). P450 can encode key enzymes of plant insect-resistant substances such as allene oxide synthase (AOS), hydroperoxide dehydratase (EC.4.2.1.92), hydrogen peroxidase, and divinyl ether synthase (Lortzing and Steppuhn 2016). In london planetree, P450 genes have not yet been cloned and reported.

We analyzed and identified the P450 genes in london planetree using bioinformatics analysis. Our goal was to use phylogenetic studies, expression characteristics, evolutionary analysis, and structural analysis of key genes to increase knowledge of the roles of P450 genes in london planetree.

Materials And Methods

Identification of CYP450 genes in london planetree

We used two methods to identify the cytochrome P450 genes in london planetree. First, we obtained the hidden Markov model (HMM) file (p450.hmm file; Pfam number: PF00067) for the cytochrome P450 family (version 26.0; http://pfam.xfam.org/(Punta et al. 2012). The software platform HMMER3 (Eddy 2011) was used to search the p450.hmm against the amino acid database of transcriptomes of london planetree, which we obtained from previous research (NCBI BioProject ID number PRJNA484863 and accession numbers SRR8631807 to SRR8631818) (F. Li et al. 2019; Fengqi Li et al. 2019). Second, we used BLASTP to identify the london planetree P450 genes (Altschul et al. 1990). All known Arabidopsis P450 gene sequences were obtained from http://p450.kvl.dk/(Paquette et al. 2009). These sequences were used as queries to search against the london planetree hypothetic protein databases using BLASTP with an e-value cut-off of 1E-5. Other options were set at default values. All identified candidate genes were submitted to NCBI Conserved Domain Search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) for conservative domain analysis, and only reserved genes with complete P450-specific domains were retained for downstream analysis. All the predicted P450 genes were assigned corresponding names according to the standardized P450 nomenclature system (Nelson et al. 1996). The functions of identified 96 P450 genes were annotated by the blast2go(Conesa et al. 2005).

Phylogenetic and Evolutionary analysis of london planetree P450 genes

All london planetree CYP450 sequences were selected for phylogenetic analysis. Multiple sequence alignment was performed using mafft software (-maxiterate 1000, using L-INS-I algorithm)(Standley 2013). A maximum likelihood (ML) phylogenetic tree was constructed using the iqtree web server (Lam-

Tung et al. 2015) (http://iqtree.cibiv.univie.ac.at). The substitution model employed was LG + F + I + G4. The ML tree was built with 1,000 ultrafast bootstrap [27] replications.

To analyze the evolutionary relationship of the P450 gene family between the london planetree and the dicotyledon model plant *Arabidipsis thaliana*, we first identified the orthologous genes between the two species using the reciprocal best hit method (Tatusov et al. 1997). Proteins and CDS sequences of P450s in *A. thaliana* were downloaded from http://p450.kvl.dk/. Then the orthologous pairwise synonymous (Ks) and non-synonymous (Ka) were estimated by the YN method of KaKs_calculator (Yu 2010).

Expression Pattern and WGCNA Analysis of P450 genes in london planetree

The 12 samples of transcriptomes from london planetree were obtained previously (F. Li et al. 2019). The 12 samples included four treatments, each treatment with three biological repetitions. The 12 samples of transcriptomes from london planetree used previously were analyzed. The four treatments were: leaves damaged by sycamore lace bug for 24 h, leaves not damaged by sycamore lace bug for 24 h (control), leaves damaged by sycamore lace bug for 48 h, and leaves not damaged by sycamore lace bug for 48 h (control). The expression of each gene was calculated by mapping the clean data of each sample to the transcriptome with the RSEM software package (Dewey and Li 2011). The gene expression difference was analyzed and compared with DEGseq2 (Likun et al. 2010). Selection of log2 (fold change) > 1 or log2 (fold change) < 1, and the statistical significance (*P* adjust value < 0.05) is significantly different genes.

To identify co-expressed P450 gene modules, TPM values were used as standardized gene expression levels, and this study included 12 samples. The gene of filtering variance was 0, as well as genes that had more than 10% of the deletion sample, using R (V.3.6.1) software (https://cran.r-project.org/bin/windows/base/old/) WGCNA (V.1.70-3) package (Peter et al. 2008) builds a weighted gene to circulate co-expressed network and divide the module. The PickSoftThreshold command was used to determine the best soft threshold and the corresponding average connectivity. We used the automatic network building function blockwise Modules (Langfelder and Horvath, 2008) to build a co-expression network. We set the parameters minmodule size = 10, maxBlockSize = 300, and other parameters according to the default settings. The WGCNA co-expression network was visualized by Cytoscape (v.3.6.1)(Shannon and P. 2003) software (Su et al., 2014). The hub genes were identified by CytoHubba (Chin et al. 2014) plug-in.

qPCR analysis

P450 gene expression was verified by selecting leaf samples and corresponding controls 24 h and 48 h after the leaves of london planetree were damaged by sycamore lace bug. The treatment method was described previously(F. Li et al. 2019). We selected healthy london planetree leaves for testing, separated them from the base of the petiole with sterile surgical blade, and inserted them into a glass finger tube containing 10 ml of sterile water. The nozzle was sealed with degreasing cotton and placed into a 2 L glass beaker. Each leaf was infested with 100 newly emerged adults of sycamore lace bug for feeding

hazard induction. The experiment was conducted in an artificial climate box at 25 ± 1° C, relative humidity of 70 ± 5% and a 16:8 h (L:D) photoperiod. After the conclusion of the experiment, insects on the leaves were removed with a brush. RNA was extracted with Trizol. The integrity and the purity of total RNA were verified using an Agilent Bioanalyzer 2100 (Agilent Technologies, CA, USA). RNA samples with a minimum RNA integration number of 6 were selected for reverse transcription. cDNA was synthesized with a prime script RT reagent kit (Perfect Real Time) (TaKaRa Bio Inc. Dalian, China). The synthesized cDNA was stored at -20°C. Using IDT Primer Quest web tool (http://www.idtdna.com/primerquest/Home/Index) to design qRT-PCR primers, GoTaq 2-step qRT-PCR system (Promega, Madison, WI, USA) was used. The qRT-PCR was then run on an ABI Prism17500 (Applied Biosystems, Foster City, California, USA). The reaction conditions were: 95°C initial denaturation 10 S, 95°C denaturation 30 S, 60°C annealed 30 S, 40 cycles. Actin (T78316_c0_g1) and alpha tubulin (T106409_c0_g1) were selected as reference genes, and the primer sequences are shown in Table S2. The inner reference gene was used to initiate the normalization of gene expression in qRT-PCR analysis, and the relative expression level of each gene was obtained by comparing 2^{-DDCt}(Livak and Schmittgen 2001).

Cloning, homology modeling, and molecular docking of AOS gene

Based on the ORF sequence of T107931_c2_g1, the primer of the ORF region of the gene was designed using Primer Premier 5, and the F primers was: ATGGGTTCTTCTTCGAGAAAAATCA, the R primer was: TCATATGTCTTCTCCCGACGCC. PCR amplification was conducted using high-fidelity PrimeSTAR HS DNA Polymerase, and the PCR product was detected by 1% agarose gel electrophoresis. We recovered a segment of the PMD18-T cloned carrier and sequencing was performed. According to the T107931_c2_g1 protein sequence, the pdb protein crystal structure database was homologous. We selected the homologous template for homologous modeling, homologous modeling and retained the Protoporphyrin Ix containing Fe amount of the original template when modeling. Three models were built and we selected the smallest discrete optimized potential energy (DOPE) and probability Density Function (PDF) value for the best model for subsequent analysis (Ali and Blundell 1993). The Structural Analysis and Verification Server (HTTPS: / /saves.mbi.ucla.edu/) the online tools PROCHECK(Laskowski et al. 1993), ERRAT(Colovos and Yeates 1993), and VERIFY 3D (Bowie et al. 1991) were used to evaluate the modeling result. The preparative protein and the preparation of the small molecule modules were treated separately. The molecules of the london planetree AOS protein and (9z, 11e, 13s) -13-HydroxyoctAdeca-9, 11-Dienoic Acid were minimized by the implementation of the CHARMM force field. Molecular docking analysis with the CDOCKER module (Gagnon et al. 2016), “Top Hits” was set to “10”; “Pose Cluster Radius” was set to “0.1”, and the others used the default parameters. We selected the area from Protoporphyrin Ix containing Fe3 Å as the active pocket center. All calculations are carried out at the Discovery Studio (Accelrys, San Diego, CA, USA) Version 2.1.

Results

Identification and classification of london planetree P450 genes

P450 genes belong to a complex gene superfamily. According to sequence homology, this superfamily can be divided into many gene families and subfamilies (Nelson et al. 1996). The P450 genes were identified from the london planetree transcriptomes using an iterative process. Consequently, a total of 96 assumed P450 genes were identified (Supplementary data 1 and 2). They all have complete P450 cytochrome motifs. Each P450 gene was assigned a name according to the standard system of P450 nomenclature described by Nelson (Nelson et al. 1996). Table 1 shows the 96 P450 genes divided into 8 clans consisting of 41 families and 61 subfamilies. Among them, the CYP71 clan, which represents the whole set of A type P450 genes, contains 44 genes belonging to 16 families (CYP71, CYP73, CYP75-CYP79, CYP81-CYP82, CYP84, CYP89, CYP92, CYP98, CYP701, CYP706, CYP736). The remaining genes were of the non-A type and belong to seven CYP clans (CYP51, 72, 74, 85, 86, 97, and 727) and 25 families (CYP51, CYP72, CYP74, CYP85-88, CYP90, CYP94, CYP96-97, CYP704, CYP707, CYP711, CYP714-716, CYP720-722, CYP727, CYP729, CYP734, CYP749, and CYP865).

Table 1
Characterization of the P450 Genes identified in london planetree.

Gene ID	Type	Clan	Family	Subfamily	AA length
T106607_c0_g1	A	71	71	CYP71AS26	507
T112746_c1_g3	A	71	71	CYP71AQ23	524
T27211_c0_g1	A	71	71	CYP71AN82	530
T64098_c0_g1	A	71	71	CYP71AS25	478
T96249_c0_g1	A	71	71	CYP71AS27	503
T104955_c0_g1	A	71	73	CYP73A321	534
T114704_c5_g1	A	71	73	CYP73A320	506
T70192_c0_g1	A	71	73	CYP73A319	421
T110750_c6_g1	A	71	75	CYP75A136	517
T111990_c2_g1	A	71	76	CYP76T42	528
T99173_c0_g1	A	71	76	CYP76A135	510
T111776_c0_g2	A	71	76	CYP76AG10	455
T197183_c0_g1	A	71	76	CYP76B132	477
T89372_c0_g1	A	71	76	CYP76AU6	510
T91529_c0_g1	A	71	76	CYP76T43	505
T94838_c2_g1	A	71	77	CYP77A91	513
T100522_c1_g1	A	71	78	CYP78A469	526
T107496_c3_g1	A	71	78	CYP78A468	439
T95772_c0_g1	A	71	78	CYP78A470	518
T114122_c11_g2	A	71	79	CYP79A209	427
T94418_c0_g1	A	71	79	CYP79A208	554
T114251_c2_g1	A	71	81	CYP81B213	510
T109002_c0_g1	A	71	81	CYP81BG48	507
T107370_c3_g1	A	71	82	CYP82C145	531
T107630_c2_g2	A	71	82	CYP82S29	562
T112128_c5_g1	A	71	82	CYP82C146	536
T96961_c1_g1	A	71	82	CYP82C147	447

Gene ID	Type	Clan	Family	Subfamily	AA length
T93893_c2_g2	A	71	84	CYP84A170	488
T106029_c0_g1	A	71	89	CYP89A333	365
T108794_c0_g1	A	71	89	CYP89A332	513
T106738_c2_g1	A	71	92	CYP92A252	517
T106738_c2_g4	A	71	92	CYP92A253	469
T131447_c0_g1	A	71	92	CYP92A251	515
T109296_c3_g1	A	71	98	CYP98A182	552
T89645_c0_g1	A	71	701	CYP701A129	511
T102543_c1_g1	A	71	706	CYP706C128	462
T105878_c2_g1	A	71	706	CYP706C129	526
T108181_c2_g1	A	71	706	CYP706C127	533
T110558_c0_g1	A	71	706	CYP706C124	438
T112194_c3_g2	A	71	75	CYP75B193	523
T113094_c1_g1	A	71	706	CYP706C126	572
T82913_c0_g1	A	71	706	CYP706C125	454
T82913_c0_g2	A	71	706	CYP706C130	529
T113051_c0_g1	A	71	736	CYP736A361	511
T111096_c0_g1	non-A	51	51	CYP51G1	447
T106385_c7_g1	non-A	72	72	CYP72A951	515
T114931_c4_g3	non-A	72	72	CYP72A952	401
T96587_c0_g1	non-A	72	72	CYP72D43	530
T98958_c0_g1	non-A	72	72	CYP72D44	520
T105299_c1_g3	non-A	727	727	CYP727B26	574
T106140_c0_g2	non-A	74	74	CYP74B58	481
T107931_c2_g1	non-A	74	74	CYP74A187	471
T108368_c2_g1	non-A	74	74	CYP74A186	503
T108368_c2_g2	non-A	74	74	CYP74A188	528
T93632_c0_g1	non-A	74	74	CYP74A185	505

Gene ID	Type	Clan	Family	Subfamily	AA length
T110273_c2_g1	non-A	85	85	CYP85A1	434
T100547_c0_g1	non-A	86	86	CYP86A259	520
T100547_c0_g2	non-A	86	86	CYP86A260	532
T112826_c0_g1	non-A	86	86	CYP86A261	561
T12699_c0_g1	non-A	86	86	CYP86C31	517
T113141_c1_g1	non-A	86	86	CYP86B101	570
T97428_c0_g2	non-A	86	96	CYP96A271	520
T105851_c0_g1	non-A	85	87	CYP87B52	509
T112210_c4_g2	non-A	85	87	CYP87B53	473
T102814_c1_g1	non-A	85	88	CYP88A155	370
T107060_c0_g1	non-A	85	729	CYP729A66	485
T103609_c1_g1	non-A	85	720	CYP720A1	494
T106945_c6_g1	non-A	85	90	CYP90B102	342
T105131_c0_g1	non-A	85	90	CYP90C52	486
T104689_c1_g1	non-A	85	90	CYP90D77	473
T107552_c0_g1	non-A	86	94	CYP94A139	418
T107552_c0_g3	non-A	86	94	CYP94A140	385
T84507_c0_g1	non-A	86	94	CYP94A141	512
T103954_c0_g1	non-A	86	94	CYP94C184	507
T104245_c0_g1	non-A	86	94	CYP94F27	460
T106684_c3_g1	non-A	86	94	CYP94D183	507
T110762_c4_g1	non-A	97	97	CYP97B153	580
T98930_c0_g1	non-A	97	97	CYP97C141	572
T110318_c0_g1	non-A	86	704	CYP704A300	536
T105507_c2_g2	non-A	85	707	CYP707A299	389
T107927_c2_g1	non-A	85	707	CYP707A298	390
T111367_c4_g1	non-A	85	707	CYP707A300	469
T106000_c0_g1	non-A	72	721	CYP721A105	517

Gene ID	Type	Clan	Family	Subfamily	AA length
T93659_c0_g1	non-A	72	865	CYP865A3	513
T109414_c0_g2	non-A	74	711	CYP711A211	549
T65915_c0_g1	non-A	72	714	CYP714G31	513
T96793_c0_g1	non-A	72	714	CYP714G32	515
T100178_c1_g1	non-A	72	715	CYP715A87	536
T104170_c4_g1	non-A	85	716	CYP716A358	539
T104170_c5_g1	non-A	85	716	CYP716A138	497
T104170_c6_g1	non-A	85	716	CYP716A139	483
T93105_c0_g1	non-A	85	716	CYP716A359	483
T102826_c0_g1	non-A	85	722	CYP722A1	508
T102256_c0_g1	non-A	72	734	CYP734A115	537
T109507_c0_g1	non-A	72	749	CYP749A371	531
T82736_c0_g1	non-A	72	734	CYP734A116	399

Phylogenetic and evolutionary analysis of predicted P450s in london planetree

The 96 london planetree P450 proteins were used to construct an ML tree (Fig. 1). The phylogenetic tree shows that 45.45% (45 genes) of the 96 CYP450s are A-type and belong to 16 families. The remaining 44.44% of CYP450s genes (44) are non-A type including seven clans and 22 families. There were eight clans in london planetree CYP450s. Four clans consist of multiple families, CYP71, CYP72, CYP74, CYP85, CYP86, and CYP88. The other five clans contain only one family each, CYP51, CYP97.

Using the reciprocal best hit method, we identified 48 P450 orthologous genes between london planetree and *A. thaliana* (Table 2). Among the 48 homologous genes, 31 were non-A type. Then, we used maximum likelihood analyses of ka and ks to estimate P450 genes that evolved under positive selection in the 48 P450 orthologous pairwise comparisons from these two species. Of these genes, all P450 orthologous genes have ka/ks < 0.17 (Table 2). This finding suggests a strong negative selection of P450 genes between london planetree and *A. thaliana*.

Table 2
Ka, Ks, and Ka/Ks values for P450 orthologous gene pairs between london planetree and *A. thaliana*.

P. acerifolia	*A. thaliana*	Ka	Ks	Ka/Ks
T100178_c1_g1	AtCYP715A1	0.312395	3.37001	0.092698
T100522_c1_g1	AtCYP78A9	0.215629	3.29902	0.065362
T100547_c0_g2	AtCYP86A1	0.150938	3.62665	0.041619
T102256_c0_g1	AtCYP734A1	0.164276	3.74829	0.043827
T102814_c1_g1	AtCYP88A3	0.218112	3.6329	0.060038
T102826_c0_g1	AtCYP722A1	0.280709	3.624	0.077458
T103609_c1_g1	AtCYP720A1	0.186885	3.92075	0.047666
T103954_c0_g1	AtCYP94C1	0.243146	3.71293	0.065486
T104170_c5_g1	AtCYP716A1	0.280525	3.57183	0.078538
T104689_c1_g1	AtCYP90D1	0.241459	3.92495	0.061519
T105131_c0_g1	AtCYP90C1	0.256305	2.67187	0.095927
T106000_c0_g1	AtCYP721A1	0.327401	3.30349	0.099108
T106140_c0_g2	AtCYP74B2	0.273318	3.48284	0.078476
T106385_c7_g1	AtCYP72A15	0.29698	3.1155	0.095324
T106607_c0_g1	AtCYP71B10	0.360631	3.262	0.110555
T106684_c3_g1	AtCYP94D2	0.286939	3.6059	0.079575
T106945_c6_g1	AtCYP90B1	0.148174	3.55165	0.04172
T107370_c3_g1	AtCYP82C4	0.313402	3.25326	0.096335
T107496_c3_g1	AtCYP78A10	0.244669	3.67622	0.066555
T107927_c2_g1	AtCYP707A4	0.179418	3.73593	0.048025
T108181_c2_g1	AtCYP706A4	0.326058	3.03185	0.107544
T108368_c2_g2	AtCYP74A	0.243616	3.49106	0.069783
T108794_c0_g1	AtCYP89A6	0.340723	3.08638	0.110395
T109002_c0_g1	AtCYP81D2	0.339567	3.05808	0.111039
T109296_c3_g1	AtCYP98A3	0.135201	3.7956	0.03562
T109414_c0_g2	AtCYP711A1	0.20527	3.66237	0.056048

P. acerifolia	*A. thaliana*	Ka	Ks	Ka/Ks
T110273_c2_g1	AtCYP85A2	0.208517	3.84729	0.054199
T110318_c0_g1	AtCYP704A1	0.282033	3.61987	0.077912
T110762_c4_g1	AtCYP97B3	0.141007	2.17206	0.064918
T111096_c0_g1	AtCYP51G1	0.083398	4.01241	0.020785
T111367_c4_g1	AtCYP707A3	0.150369	3.95456	0.038024
T112194_c3_g2	AtCYP75B1	0.223386	3.25441	0.068641
T112210_c4_g2	AtCYP87A2	0.37237	3.1471	0.118322
T112746_c1_g3	AtCYP71A26	0.365246	3.05039	0.119737
T112826_c0_g1	AtCYP86A2	0.171169	3.62077	0.047274
T113141_c1_g1	AtCYP86B1	0.218138	3.79198	0.057526
T114704_c5_g1	AtCYP73A5	0.087907	3.9213	0.022418
T12699_c0_g1	AtCYP86C1	0.279705	3.56262	0.078511
T84507_c0_g1	AtCYP94B1	0.466905	2.79813	0.166863
T89645_c0_g1	AtCYP701A3	0.265476	2.53937	0.104544
T91529_c0_g1	AtCYP76C4	0.435223	2.70903	0.160657
T93893_c2_g2	AtCYP84A1	0.150196	4.06827	0.036919
T94418_c0_g1	AtCYP79A2	0.310712	3.50924	0.088541
T94838_c2_g1	AtCYP77A4	0.198313	3.71916	0.053322
T96793_c0_g1	AtCYP714A1	0.435318	2.98695	0.14574
T97428_c0_g2	AtCYP96A1	0.430124	2.65867	0.161781
T98930_c0_g1	AtCYP97C1	0.126936	3.9521	0.032119
T99173_c0_g1	AtCYP76G1	0.417738	2.94527	0.141833

Gene expression profiles and co-expression network analysis of london planetree P450 genes

RNA sequencing was used to analyze the gene expression partners of all 96 CYP450s. The results reveal the expression pattern of the P450 genes involved in the response of london planetree leaves to the feeding stress of sycamore lace bug. After 24 h of feeding damage, 26 P450 genes were significantly up-regulated and 4 P450 genes (T96793_c0_g1, T108181_c2_g1, T110558_c0_g1, T102543_c1_g1) were significantly down-regulated (Table S3). After 48 h of damage, 12 P450 genes were significantly up-

regulated and 4 P450 genes (T96587_c0_g1, T105507_c2_g2, T109002_c0_g1, T108368_c2_g2) were significantly down regulated. Among them, 10 P450 genes were simultaneously significantly up-regulated after 24 h and 48 h of feeding damage (Table S3 and Fig. 2 and Figure S1) as common defense genes. No P450 gene that were simultaneously significantly down-regulated were identified at both time point.

The samples were clustered according to the gene expression level. The clustering results based on average distance and the hierarchical clustering algorithm show that the materials used for sequencing have high repeatability in each process and the gene expression patterns in the same process are similar and clustered together (Fig. 2). We calculated the soft valve value (β) and finally the β value set as 11. The weighted gene co-expression network was successfully constructed by the WGCNA method, and the network was divided into four main modules (Figure S3, Table S4). The number of genes in the co-expression module ranged from 14–40, and a total of 97 genes were contained. The number of genes containing the turquoise module was greatest (up to 40) while at least 13 genes were present in the yellow module. The yellow module corresponded to the defensive mode after the sycamore lace bug damage. (Figure S4). Thus, the module may participate in anti-insect resistance. Further analysis of the yellow module identified three hub genes (CYP74A187 (T107931_c2_g1), CYP94A140 (T107552_c0_g3), and CYP79A208 (T94418_c0_g1) (Fig. 3). Among these three genes, T107931_c2_g1 are key oxide synthase encoding genes of JA synthesis pathways, T107552_c0_g3 is P450-dependent fatty omega-hydroxylase, which is a key enzyme that forms hydroxy-fatty. T94418_c0_g1 is a gene encoding phenylalanine N-monooxygenase.

Qpcr Verification Of Key P450 Genes Of London Planetree Against Sycamore Lace Bug

We chose the three hub genes identified above, and 6 common up-regulated genes. A total of 9 P450 genes were verified by qPCR expression. The qPCR result was consistent with the result of the quantitative analysis of the transcription group (Fig. 4). Among them, six genes including T107931_c2_g1 (CYP74A187), T107552_c0_g3 (CYP94A140), T94418_c0_g1 (CYP79A208), T82913_c0_g2 (CYP706C130), T107630_c2_g2 (CYP82S29), and T111990_c2_g1 (CYP76T42) were expressed after sycamore lace bug feeding damage. And, the level of expression of these six genes at 48 h was higher than that at 24 h. Four genes, including T91529_c0_g1 (CYP76T43), T106140_c0_g2 (CYP74B58), T107630_c2_g2 (CYP82S29) and T108368_c2_g1 (CYP74A186) were expressed at 24 and 48 h after sycamore lace bug damage but their expression at 48 h was less than that at 24 h. In these genes, the qPCR results showed that T107931_c2_g1 (CYP74A187), T111990_c2_g1 (CYP76T42), T91529_c0_g1 (CYP76T43), T107630_c2_g2 (CYP82S29), and T108368_c2_g1 (CYP74A186) have dominated and may play a relatively more important role in london planetree defense against sycamore lace bug.

Cloning of AOS genes, homologous modeling and molecular docking

From the transcriptome expression quantity of the above genes and the results of qPCR analysis, we found that T107931_c2_g1 (CYP74A187) was significantly up-regulated after 24 h and 48 h. This gene is the key enzyme gene of JA synthesis pathway (Ruan et al. 2019). The induced expression of this gene suggests that the JA pathway is involved in the defense of london planetree against sycamore lace bug. In view of this, we further cloned this AOS gene (T107931_c2_g1), and conducted homologous modeling and molecular docking analysis of the AOS gene. Cloning and sequencing showed that the full-length sequence of the AOS gene of london planetree was 1485 bp and it encoded 494 amino acids (NCBI Genbank number OM638588). Three templates were further selected to model the homology of london planetree AOS protein, The selected PDB templates included CYP74A (PDB ID: 2RCH) and CYP74A (PDB ID: 2RCM) of *A. thaliana* (Lee et al. 2008)and CYP74A of *Parthenium argentatum* (PDB ID: 3DAM) (Li et al. 2008). These templates had protein level homology of 63.14%, 62.92%, and 61.06%. Three models were constructed of which the DOPE score of model 2 (-53553.414063) was the lowest after further analysis with Profiles-3D, the verify score of the model was 164.53, the expected high score was 215.099 and the expected low score was 96.7944. The Ramachandran diagram of Procheck evaluation results (Figure S5) show that the modeling protein 92 6% of the residues were located in the optimal region, 5.9% of the residues fall in other suitable areas, and no residues fall in reluctantly permitted areas and unreasonable areas. These results reached a high-quality modeling standard (> 90%) and showed that the results obtained by modeling are reasonable. Molecular docking analysis showed that (9z, 11e, 13s) − 13-hydroxyadeca-9,11-dienoic acid could be bound by london planetree AOS, and the highest binding energy was − 35.5538 kcal mol-1 highest interaction energy was − 45.9477 kcal mol-1. Molecular interaction analysis showed that there were alkyl and PI alkyl forces between four amino acids CYS140, LYS154, PHE157, and PRO460 and small molecules, amino acids TRY380 and ASN445 were conventional hydrogen bonds for AOS and small molecules. There was an attractive charge between LYS119 and small molecules, and there was a carbon-hydrogen bond between GLN379 and small molecules (Fig. 5). These eight amino acids are the key amino acid residues of the interaction between london planetree AOS and (9z, 11e, 13s) − 13-hydroxyactadeca-9,11-dienoic acid.

Discussion

The extensive worldwide planting of london planetree is closely related to its excellent growth characteristics, pollution treatment ability, and strong stress resistance. The cytochrome P450 gene superfamily is one of the largest enzyme families supporting plant metabolism. This explains its importance in defense responses, metabolite synthesis, and signal transduction. The diversity of P450 species and the universality of substrates indicate diverse functions. P450 genes play important roles in almost all aspects of plant growth and development. The cytochrome P450 gene is closely related to plant growth and development, detoxification metabolism, and stress tolerance. The current understanding of this important tree species and its P450 genes is limited; this constrains our understanding of the development, metabolism, and stress tolerance of london planetree. This study systematically identified the P450 gene family of london planetree and analyzed its evolution and

expression characteristics. These data and results improved our understanding of the biological and physiological characteristics of london planetree.

In this study, we identified 96 P450 genes of london planetree that are distributed in 41 families and 61 subfamilies. There are few reports of P450 genes in london planetree which is inconsistent with results on other dicotyledonous plants. This may be related to the lack of a genome sequence for london planetree. Among dicotyledons, the genome sequencing of *A. thaliana*, tomato, alfalfa, and other plants has been completed, and many P450 genes have been identified in these species. However, research on the london planetree genome is far behind other dicotyledons so new P450 will likely be discovered in the future. As an ancient species, london planetree may have more abundant P450 family proteins to be identified.

The majority (51) of the 96 P450 genes of london planetree have no direct orthologous genes in *A. thaliana*, indicating that P450 family may have a high degree of differentiation in dicotyledons. Among all of the reported P450 genome families, P450 has differentiated into many gene families unique to plants.

This study identified 44 A-Type gene members and 52 Non-A-Type gene members. Among the pairing of *A. thaliana*, Non-A-Type gene members account for high proportions (64.58%). This is because Non-A-Type gene members are universally conserved in the plant P450 gene superfamily, which is associated with gene conservation and gene functionality. In plants, the Non-A-Type CYP51 families and CYP85 subfamilies mainly participate in the synthesis of solids and steroids. CYP86 families mainly participate in the metabolism of fatty acids. The CYP97 family mainly participates in the biological synthesis of terpenoids and the CYP72 family mainly participates in the metabolism of plant hormones. These features are required for plant growth and metabolism. The CYP71 family of P450 has a diversity of functions to deal with plant-specific needs (Schuler 2011).

A ka/ks value of less than 1 is often a sign of negative selection or a purification selection and a value greater than 1 is considered to be a sign of positive selection (Li et al. 2015; Yang and Bielawski 2000). We used positive selection analysis and discovered that the ka/ks values in our study in all orthologous genes were less than 0.17. This indicates that some P450 genes shared between london planetree and *A. thaliana* are highly conserved and are operating under purification selection, suggesting that they have important functions and are strong functional constraints. This also indirectly illustrates the importance of P450 genes in the life cycles of these two plants.

Sycamore lace bug is an invasive species and one of the most important pests of london planetree (D. Li et al. 2019; Feng-Qi et al. 2018). In this study, we identified nine P450 genes involved in the defense of london planetree against sycamore lace bug. These genes may be involved in the synthesis and metabolism of hormones and toxic substances related to insect resistance. These genes provide a basis for understanding the interaction between london planetree and sycamore lace bug and provide a reference for improving the insect resistance of london planetree in future (Schuler 2011).

T82913_c0_g2 (CYP706C130) is a flavonoid 3 '- monooxygenase, and a key enzyme in the flavonoid metabolic pathway. Flavonoid metabolites are closely related to the insect resistance of plants. T107630_c2_g2 (CYP82S29) belongs to the CYP82 family, which is involved in the biosynthesis of plant terpene ketone series DMNT (Li et al. 2018; Lee et al. 2010). Leaves of london planetree produced DMNT after being induced by sycamore lace bug feeding (Feng-Qi et al. 2017). T107630_c2_g2 (CYP82S29) may be a key gene for the synthesis of DMNT. DMNT often attracts natural enemies and plays an indirect role in plant anti-insect responses. The specific function of this gene could be confirmed by a series of functional verification such as yeast expression in future.

Jasmonic acid is important in plant resistance to insect attack (Lortzing and Steppuhn 2016). The key insect resistance genes identified in this study were one HPL gene and two AOS genes. Both of these genes are involved in the jasmonic acid synthesis pathway. Among them, T106140_c0_g2 (CYP74B58) is a fatty acid hydroperoxide lyase (HPL) gene, which is a key gene in the jasmonic acid synthesis pathway. Allene oxide synthase (AOS) is a member of the P450 superfamily that helps regulate many plant biological processes by controlling JA biosynthesis [39]. *Spodoptera litura* (Fabricius) feeding damage can significantly induce the rice jasmonic acid synthesis pathway expressing the lipoxygenase and propylene oxide synthase genes in the insect-induced defense system (Tao et al. 2003). Injury can induce expression of the AOS gene in tomato leaves, and the level of AOS mRNA can be 5 to 9 times that of the corresponding control (Sivasankar and Rothstein 2000). Enhancement of AOS gene expression improves plant adaptation to the physical environment and helps plants resist pest attack. In this study, two significantly up-regulated AOS genes, T107931_c2_g1 (CYP74A187) and T108181_c2_g1 (CYP706A), were identified, and their expression was significantly up-regulated at 24 h and 48 h after leaf damage caused by sycamore lace bug. T107931_c2_g1 (CYP74A187) is the hub gene of the P450 gene expression module. The expressions of HPL and AOS genes in this study suggested that the jasmonic acid pathway is important in the interaction between london planetree and sycamore lace bug. Our structural modeling and molecular results provide clues for studying the specific function of the london planetree AOS gene.

Declarations

Author Contributions Conceptualization, F.L.; methodology, C.W.; software, C.W.; formal analysis, A.T.; writing—original draft preparation, C.W.; writing—review and editing, Y.D.; funding acquisition, A.T. All authors have read and agreed to the published version of the manuscript.

Funding This study was supported by National Natural Science Foundation of China, grant numbers 3010010259, and the projects of Beijing Municipal Natural Science Foundation, grant number 6202005.

Acknowledgments: We thank Dr. David R. Nelson for naming the london planetree P450s.

Conflicts of Interest: The authors declare no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Ali A, Blundell TL (1993) Comparative Protein Modelling by Satisfaction of Spatial Restraints. J Mol Biol 234(3):779–815
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410. doi:10.1016/S0022-2836(05)80360-2
3. Bowie J, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science 253(5016):164–170
4. Chen C, Chen H, Zhang Y, Thomas HR, Xia R (2020) TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant 13:8
5. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY (2014) cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 8(S4):S11
6. Colovos C, Yeates TO (1993) Verification of protein structures: Patterns of nonbonded atomic interactions.Protein Science
7. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18):3674–3676
8. Dewey CN, Li B (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12(1):323–323
9. Eddy SR (2011) Accelerated Profile HMM Searches. PLoS Comput Biol 7(10):e1002195. doi:10.1371/journal.pcbi.1002195
10. Feng-Qi LI, Ning-Ning FU, Zhang LZ, Jiao MM, Peng LF, Yi-Hua XU, Luo C (2018) Advances in biology,chemical ecology and control of the sycamore lace bug,Corythucha ciliata (Hemiptera:Tingidae). Acta Entomologica Sinica
11. Feng-Qi LI, Shi-Yong Y, Ning-Ning FU, Cheng QU, Jia L, Ran W, Yi-Hua XU, Chen L (2017) Volatile profiles of Platanus acerifolia leaves and their behavioral effects on Corythucha ciliata(Hemiptera:Tingidae).Chinese Journal of Applied Entomology
12. Gagnon JK, Law SM, Iii C (2016) Flexible CDOCKER: Development and application of a pseudo-explicit structure-based docking method within CHARMM.Journal of Computational Chemistry
13. Halbert SE, Meeker JR (1983) The Sycamore Lace Bug, Corythucha ciliata (Say) (Hemiptera: Tingidae)1. Ann Entomol Soc Am 76(2):262–265264
14. Lam-Tung N, Schmidt HA, Arndt VH, Quang MB (2015) IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.Molecular Biology & Evolution(1):268–

274

15. Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures.Journal of Applied Crystallography**26**
16. Lee DS, Nioche P, Hamberg M, Raman CS (2008) Structural insights into the evolutionary paths of oxylipin biosynthetic enzymes. Nature 455(7211):363–368
17. Lee S, Badieyan, Somayesadat, Bevan R, David Herde, Marco, Gatz, Christiane, and Tholl. 2010. Herbivore-induced and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*
18. Li D, Dewer Y, Qu C, Li F, Luo C (2019) Metabolomics Profiling and AKR Characterization During Paurometabolous Development of Corythucha ciliata (Hemiptera: Tingidae). J Insect Sci 19(6):1–8
19. Li FQ, Fu NN, Qu C, Wang R, Xu YH, Luo C (2017) Understanding the mechanisms of dormancy in an invasive alien Sycamore lace bug, Corythucha ciliata through transcript and metabolite profiling. Rep 7(1):2631
20. Li FQ, Li W, Lin YJ, Pickett JA, Birkett MA, Wu K, Wang G, Zhou JJ (2018) Expression of lima bean terpene synthases in rice enhances recruitment of a beneficial enemy of a major rice pest. Plant, Cell & Environment
21. Li F, Wu C, Dewer Y, Li D, Luo C (2019) Changes in Gene Expression and Metabolite Profiles in Platanus acerifolia Leaves in Response to Feeding Damage Caused by Corythucha ciliata. Int J Mol Sci 20(14):3465
22. Li F, Cao D, Liu Y, Yang T, Wang G (2015) Transcriptome Sequencing of Lima Bean (Phaseolus lunatus) to Identify Putative Positive Selection in Phaseolus and Legumes. Int J Mol Sci 16(7):15172–15187
23. Li F, Wu C, Gao M, Luo C (2019) Transcriptome sequencing, molecular markers, and transcription factor discovery of Platanus acerifolia in the presence of Corythucha ciliata. Sci Data 6(1):128
24. Li L, Chang Z, Pan Z, Fu ZQ, Wang X (2008) Modes of heme binding and substrate access for cytochrome P450 CYP74A revealed by crystal structures of allene oxide synthase. Proc Natl Acad Sci USA 105(37):13883–13888
25. Likun W, Zhixing F, Xi W, Wang (2010) DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26(1):136–138
26. Livak KJ, Livak Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDCt method. Methods 25(4):402–408
27. Lortzing, and Steppuhn (2016) Jasmonate signalling in plants shapes plant-insect interaction ecology. CURR OPIN INSECT SCI
28. Mizutani M, Ohta D (2010) Diversification of P450 genes during land plant evolution. Annu Rev Plant Biol 61:291–315. doi:10.1146/annurev-arplant-042809-112305

29. Nelson DR (2011) Progress in tracing the evolutionary paths of cytochrome P450. Biochim Biophys Acta 1814(1):14–18
30. Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR et al (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics 6(1):1–42
31. Nelson D, Werck-Reichhart D (2011) A P450-centric view of plant evolution. Plant J 66(1):194–211. doi:10.1111/j.1365-313X.2011.04529.x
32. Paquette SM, Jensen K, Bak S (2009) A web-based resource for the Arabidopsis P450, cytochromes b(5), NADPH-cytochrome P450 reductases, and family 1 glycosyltransferases (. Phytochemistry 70(17–18):1940–1947. http://www.P450.kvl.dk doi:10.1016/j.phytochem.2009.08.024
33. Peter LangfelderSteve, and Horvath (2008) WGCNA: an R package for weighted correlation network analysis.BMC Bioinformatics
34. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N et al (2012) The Pfam protein families database. Nucleic Acids Res 40(Database issue):D290–301. doi:10.1093/nar/gkr1065
35. Ruan J, Zhou Y, Zhou M, Zhang K (2019) Jasmonic Acid Signaling Pathway in Plants. Int J Mol Sci 20(10):2479
36. Schuler MA (2011) P450s in plant-insect interactions. Biochim Biophys Acta 1814(1):36–45
37. Shannon, and P (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res 13(11):2498–2504
38. Sivasankar S, Sheldrick Steven J, Rothstein (2000) Expression of allene oxide synthase determines defense gene activation in tomato. Plant Physiol 122(4):1335–1342
39. Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772
40. Tao, Xu, Qiang Z, Wei, Chen G, Zhang G, He (2003) Involvement of Jasmonate-signaling pathway in the herbivore-induced rice plant defense. Chinese Science Bulletin
41. Tatusov RL, Koonin EV, Lipman DJ (1997) A Genomic Perspective on Protein Families. Science 278(5338):631–637
42. Yang Z, Bielawski JP (2000) Stat methods detecting Mol adaptation 15(12):0–503
43. Yu ZJ (2010) KaKs_Calculator 2.0: A Toolkit Incorporating Gamma-Series Methods and Sliding Window Strategies. Genomics, Proteomics & Bioinformatics
44. Zhang J, Cong G, Liu G, Li Z, Li X, Bao M (2011) Genetic alteration with variable intron/exon organization amongst five PI-homoeologous genes in Platanus acerifolia. Gene 473(2):82–91

Figures

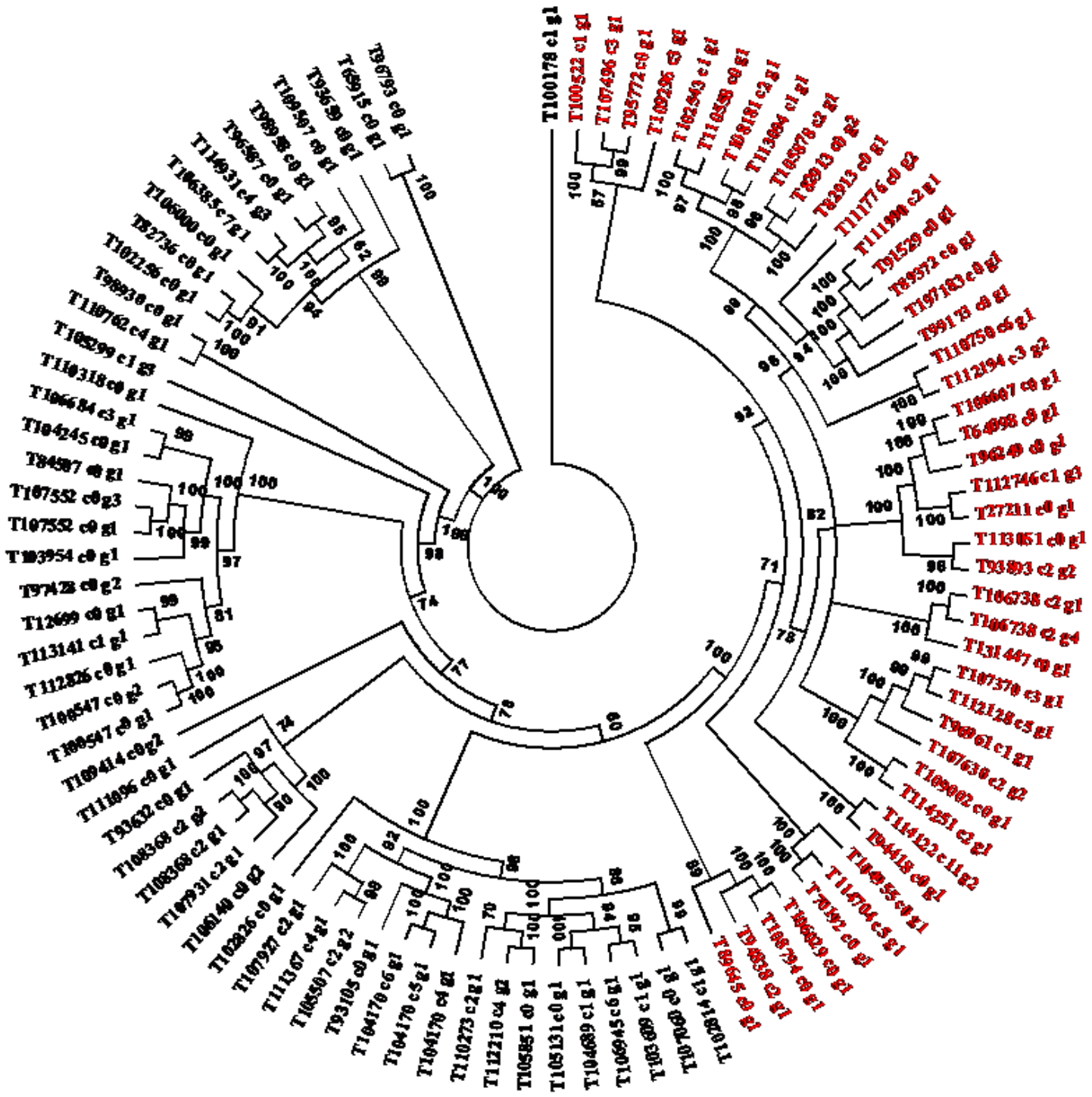

Figure 1

Phylogenetic relationships of london planetree P450 proteins. An unrooted NJ tree of P450s was constructed using the iqtree [17]. There were 1,000 bootstrap replicates. The Red represents A-type and black represents non-A type.

Figure 2

Expression of the london planetree P450 genes.

The red and green colors mean the expression levels of P450 genes from high to low, and white indicates the median expression level in the heatmap. The heatmap was generated by using TBtools software (Chen et al. 2020). The sources of the different treatment are provided on the*x*-axis. The genes were

clustered according to the Average TPM value per treatment.Expression pattern of london planetree CYP450s under feeding attack by sycamore lace bug based on the analysis of the RNA-Seq dataset. The color scale shows the expression quantity (red: high expression; green: low expression).

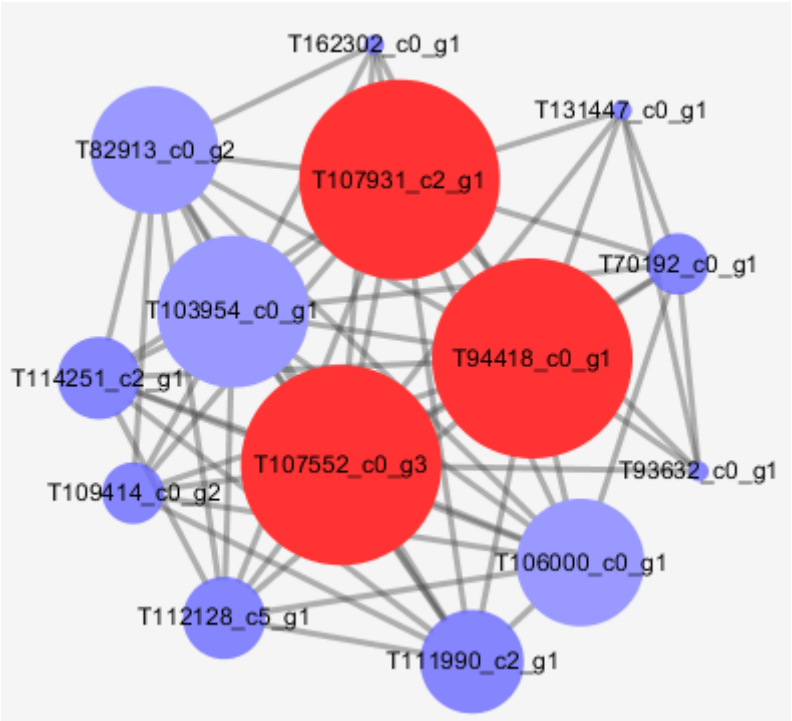

Figure 3

Co-expression network analysis of the yellow module. Edge stands for the interaction between two genes. The size of the circle represents the degree of importance in the module. The three genes marked in red are the hub genes. These genes were identified by degree method of plug-in cytoHubba in Cytoscape.

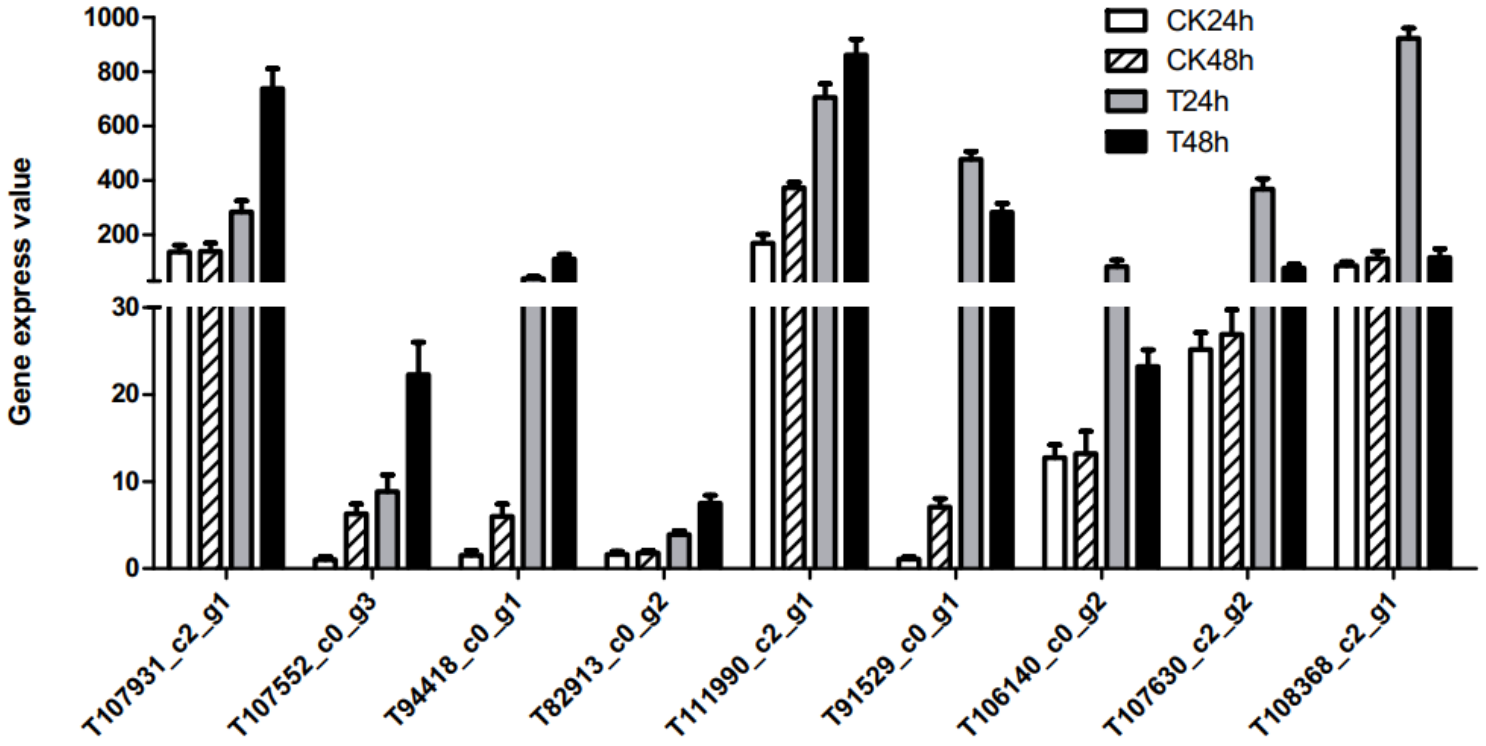

Figure 4

qRT-PCR validations results of candidate P450 in london planetree defense against sycamore lace bug . qRT-PCR were performed in triplicate. Error bars indicate the standard error of the mean.

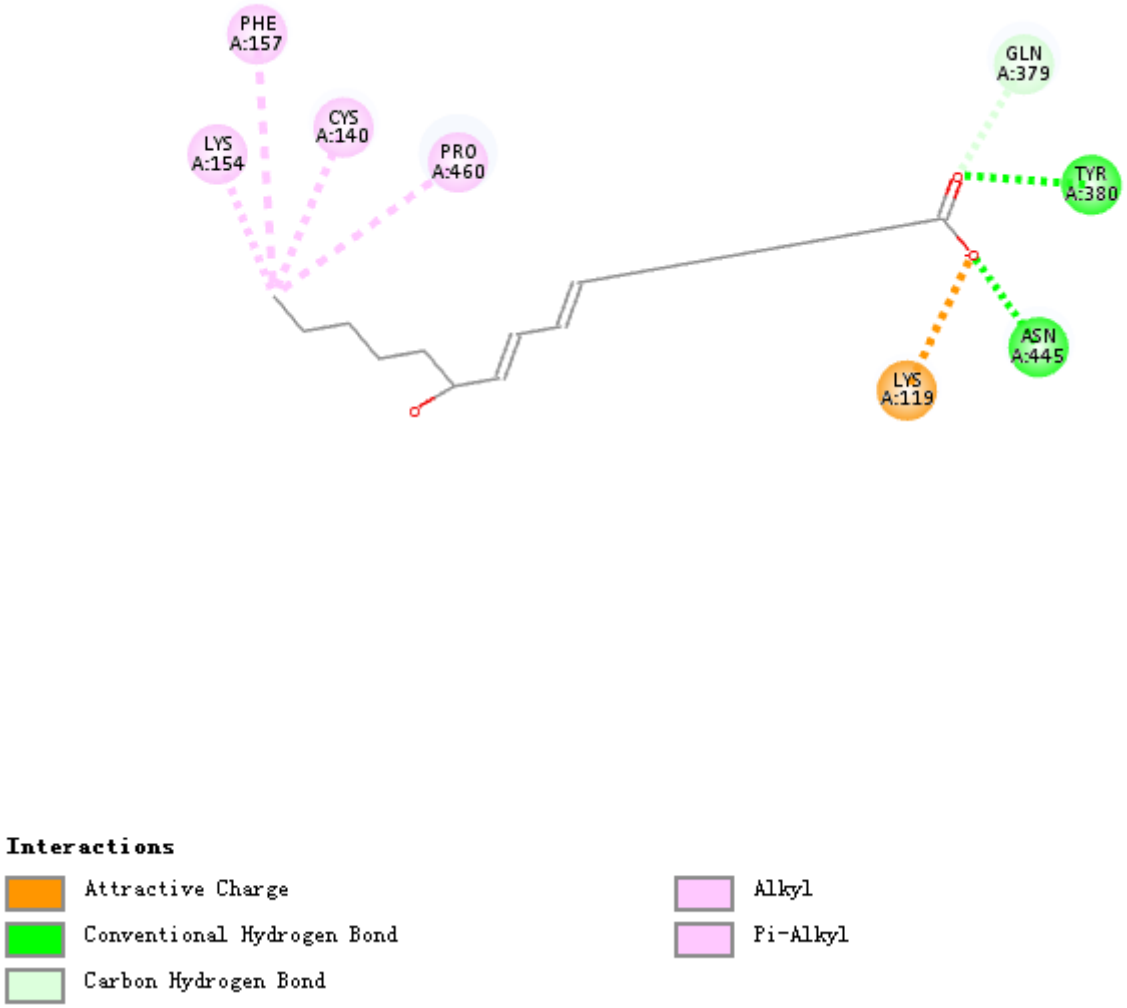

Figure 5

Diagram of the interactions of AOS and (9z, 11e, 13s) - 13-hydroxyactadeca-9,11-dienoic acid with key binding site residues.

Supplementary Files

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

- figs1.png
- figs2sampleclustering.pdf
- figs3wgcna.dendroColors.png
- figs4wgcna.yellow.express.barplot.png
- figs5.pdf
- TableS14.xlsx
- Supplementarydata1.P450cds.fasta

- Supplementarydata2.P450pep.fasta