

Article

Not peer-reviewed version

Genome-Wide Identification and Functional Analysis of WRKY Transcription Factors in Tetraploid *Camellia oleifera* Highlights a Key Regulator of Anthracnose Resistance

[Jihua Cheng](#) , Kefang Ou , Yangwen Du , Yingchun Jiang , Dezhi Jiang , Yawen Xu , [Junhua Peng](#) ^{*} , [Junyong Cheng](#) ^{*}

Posted Date: 6 January 2026

doi: 10.20944/preprints202601.0460.v1

Keywords: WRKY family; oil-tea; *Camellia oleifera*; anthracnose resistance



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Genome-Wide Identification and Functional Analysis of WRKY Transcription Factors in Tetraploid *Camellia oleifera* Highlights a Key Regulator of Anthracnose Resistance

Jihua Cheng ¹, Kefang Ou ², Yangwen Du ¹, Yingchun Jiang ³, Dezhi Jiang ¹, Yawen Xu ¹, Junhua Peng ^{4,*} and Junyong Cheng ^{1,*}

- ¹ Hubei Academy of Forestry, Wuhan 430022, Hubei, China
- ² Wuhan Institute of Landscape and Architecture, Wuhan 430022, Hubei, China
- ³ The People's Government of Futianhe Town, Macheng 438300, Hubei, China
- ⁴ Spring Valley Agriscience Co., Ltd., Jinan 250307, Shandong, China
- * Correspondence: junhuapeng@yahoo.com (J.P.); 185310075@qq.com (J.C.)

Abstract

WRKY transcription factors (TFs) are pivotal in plant stress responses, yet their roles in *Camellia oleifera*, an economically important oil crop, remain poorly understood. We identified 192 WRKY genes in the tetraploid *C. oleifera* genome, and classified them into three groups (I, II and III) based on conserved domains. Chromosomal distribution revealed uneven localization of the WRKY genes, with the highest density (25 WRKY genes) on the Chromosome 10. RNA-seq analysis on anthracnose-resistant (CL150) and susceptible (CL102) cultivars inoculated with *Colletotrichum gloeosporioides* identified 1,822 differentially expressed genes (DEGs) and 109 DEGs dependent to CL150, including 11 core DEGs shared between the cultivars. Notably, one WRKY gene (YC.08G0001620-1A, Type I) exhibited significant upregulation in CL150, suggesting its role in disease resistance. Functional enrichment linked the DEGs to oxidative stress and metabolic pathways. This study provides a comprehensive WRKY family analysis in *C. oleifera* and highlights YC.08G0001620-1A as a promising candidate for molecular breeding to enhance anthracnose resistance in this economically important oil crop.

Keywords: WRKY family; oil-tea; *Camellia oleifera*; anthracnose resistance

1. Introduction

WRKY transcription factors (TFs) represent one of the largest and most important families of regulatory proteins in plants [1]. These proteins can be classified into three major types (I, II and III) by the presence of WRKY domains (WRKYGQK) at N-termini and the configuration of zinc-finger-like motifs at C-termini [2]. Type I contains one or two WRKY domains without zinc-finger motifs, Type II possesses WRKY domains and C2H2-type zinc-finger motifs (C-X4-5-C-X22-23-H-X-H), and Type III contains WRKY domains and C2HC-type zinc-finger motifs (C-X7-C-X23-H-X-C) [3]. These domains enable WRKY proteins to specifically bind to W-box (TTGAC[T/C]) cis-elements in target gene promoters, allowing them to regulate diverse biological processes [2, 3].

Extensive researches have demonstrated that WRKY TFs play a pivotal role in modulating responses to both biotic [4-7] and abiotic stresses [8-10], as well as in developmental processes [5, 11] and hormone signaling pathways [6, 12, 13]. The investigations on molecular mechanisms of WRKY in model plants showed that WRKY proteins often function through protein-protein interactions, particularly with Valine-glutamine motifs (VQ-motif) containing proteins. For example, AtVQ10 interacts with AtWRKY8 in Arabidopsis jointly enhance resistance to *Botrytis cinerea* [14, 15]. Similar mechanisms have been observed in crops, where MdWRKY75 in apple confers resistance to *Fusarium*

solani by regulating *MdERF114* expression [5], and *VqWRKY31* in grape enhances resistance to *Erysiphe necator* through salicylic acid signaling [4]. In contrast, a more complicated module (*JrPHL8-JrWRKY4-JrSTH2L*) in walnut is involved into defense to *Colletotrichum gloeosporioides* [7].

Camellia oleifera, a native species to China, is widely used as a woody oil crop with highly valuable oil in seeds [16]. However, anthracnose disease caused by the fungal pathogen *Colletotrichum gloeosporioides* poses a significant threat to yield and quality, leading to substantial economic losses [17]. Current management strategies relying on chemical fungicides are neither sustainable nor completely effective. Recently, the release of reference genomes for different ploidy *C. oleifera* [18-20], provides an opportunity to study its WRKY family, like analyses conducted in cotton and soybean [21, 22]. These studies in other polyploid species revealed lineage-specific expansions and functional diversification of WRKY families driven by whole-genome duplication (WGD).

Despite their established importance in plant defense, the WRKY transcription factors in *C. oleifera* remains poorly characterized. To address this knowledge gap, we conducted a comprehensive study aiming to: (1) identify and classify WRKY genes in the tetraploid *C. oleifera* genome; (2) analyze their structural features and evolutionary patterns; (3) investigate their expression profiles in response to *C. gloeosporioides* infection using resistant (CL150) and susceptible (CL102) cultivars; and (4) explore their potential roles in anthracnose resistance. Our findings provide valuable insights for establishing molecular breeding strategies to enhance disease resistance in this species.

2. Materials and Methods

2.1. Plant Materials and Treatments

Two clonal varieties of *Camellia oleifera* Abel. with different anthracnose resistance, CL150 (resistant) and CL102 (susceptible), were selected for this study. Two-year-old cuttings of both varieties were cultivated in a climate chamber. Pathogen inoculation was performed following the procedure described by Yang [23]. Briefly, *Colletotrichum gloeosporioides* was cultured on potato dextrose agar (PDA) medium for 5-7 days to make 5 mm fungus cake. Prior to inoculation, young leaves were surface-sterilized with 75% ethanol followed by rinsing with sterile distilled water. Two wounds were created per leaf on either side of the midvein using a sterile needle, and each wound site was treated with 10 μ L of 1% sterile glucose solution before applying a fungal disc. Control leaves received sterile PDA plugs without fungal inoculation. The post-inoculated symptom on leaves has been shown in the previous study [16]. Three biological replicates were employed per treatment (inoculated and control) for each variety, resulting in 12 experimental samples in total (2 varieties $\times$ 2 treatments $\times$ 3 replicates). Leaf samples (500 mg) were collected at 72 hours post-inoculation (hpi) and immediately frozen in liquid nitrogen for subsequent RNA extraction.

2.2. Identification of WRKY Genes in Tetraploid *C. oleifera* Genome

The WRKY gene family was systematically identified using a dual-approach strategy combining hidden Markov model (HMM)-based searches and homology-based verification. First, the WRKY domain model (PF03106) and the protein sequences for model building (97 reviewed proteins, accessed on April 8, 2025) were obtained from InterPro database, and the complete protein sequences of tetraploid *C. oleifera* were retrieved from CNCB database (accession: GWHERT000000000). Next, the HMMER software (v3.3.2) [24] was employed to scan the *C. oleifera* proteome for potential WRKY genes using the PF03106 profile with stringent parameters (-E 1e-20, --domE 1e-20). To ensure accuracy, the candidate WRKY proteins were further validated via BLASTP analysis (NCBI BLAST+ v2.12.0) against the 97 protein sequences with the parameter -evalue 1e-20. Only sequences consistently identified by both HMMER and BLASTP were retained as high-confidence WRKY genes for subsequent analysis.

2.3. Characterization and Classification of WRKY Genes

The identified WRKY genes were further analyzed to confirm the presence of conserved domains, including WRKY motif and zinc-finger structures, using InterProScan (v5.52-86.0) [25] with default parameters. Subsequently, the WRKY genes were classified into three major types (Type I, II, and III) based on their domain architecture: Type I (lacking zinc-finger domains), Type II (containing C2H2 zinc-finger motifs), and Type III (featuring C2HC zinc-finger motifs). This classification was performed using an in-house shell script. To elucidate gene structure diversity, exon-intron organization and untranslated regions (UTRs) were determined using the tetraploid *C. oleifera* genome annotation file (GFF3 file).

2.4. Chromosomal Localization and Phylogenetic Analysis

The physical locations of all identified WRKY genes were extracted from the tetraploid *C. oleifera* genome annotation file and visualized on their respective chromosomes using the RIdiogram package (v0.2.0) [26], enabling comprehensive analysis of their distribution patterns. For phylogenetic analysis, the full-length WRKY protein sequences were aligned using ClustalW (v2.1) [27] with default parameters. A neighbor-joining phylogenetic tree was then constructed using MEGA11 software [28] with 1000 bootstrap replicates.

2.5. RNA-Seq and Differential Expression Analysis

Total RNA was extracted from leaf samples using the Trizol reagent, and RNA-seq libraries were prepared and sequenced using the Illumina NovaSeq platform. Raw sequencing reads have been deposited in the NCBI (accession number PRJNA775660). For data analysis, raw reads were first processed using fastp (v0.23.2) [29] to remove adapters and low-quality bases with default parameters. The resultant clean reads were then mapped to the tetraploid *C. oleifera* genome using BWA-MEM (v0.7.17) [30] with default parameters followed by gene expression quantification with featureCounts [31]. Differential expression analysis was conducted using DESeq2 [32], with a false discovery rate (FDR) < 0.05 and $|\log_2\text{FC}| \geq 1$ as the criteria for significantly differential expression.

2.6. Functional Annotation and Network Analysis

Due to the limited functional annotation available for the tetraploid *C. oleifera* genome, we employed a comparative genomics approach using the relatively well-annotated genome of *Camellia sinensis* var. Longjing 43 as a reference. The complete proteome and corresponding GO and KEGG annotations for *C. sinensis* were retrieved from the Tea Plant Genome Database (<https://tpia.teaplants.cn/download.html>). For functional annotation transfer, we performed protein sequence similarity searches using BLASTP (v2.12.0+) with an E-value $1e-20$. The top hits (minimum identity >40%, coverage >50%) were used to assign putative GO and KEGG terms to *C. oleifera* genes based on their orthologous relationships with *C. sinensis* genes. This homology-based annotation approach generated a preliminary functional annotation database for tetraploid *C. oleifera*. The GO and KEGG annotation analysis of differentially expressed genes (DEGs) was subsequently performed using the clusterProfiler package (v4.0.5) [33]. Statistical significance was assessed using Fisher's exact test with Benjamini-Hochberg correction (FDR < 0.01).

3. Results

3.1. Genome-Wide Identification and Chromosomal Distribution of WRKY Genes

Through a comprehensive analysis combining HMMER and BLASTP approaches, a total of 192 WRKY genes were identified in the *C. oleifera* genome. These genes unevenly distributed along the chromosomes and across the genome with exception of three chromosomes (no WRKY genes were identified on Chr09A, Chr09C and Chr13D), and these genes also unevenly distribute across the *C. oleifera* chromosomes (Figure 1, Table 1). Specifically, Chr10 (all four homologous copies: -A, -B, -C

and -D) had the highest number of WRKY genes (n=25), representing approximately 12.4% of the total WRKY genes. In contrast, Chr09 and Chr13 (-A, -B, -C and -D) showed the lowest density, harboring only 2-3 genes each (Figure 1, Table 1).

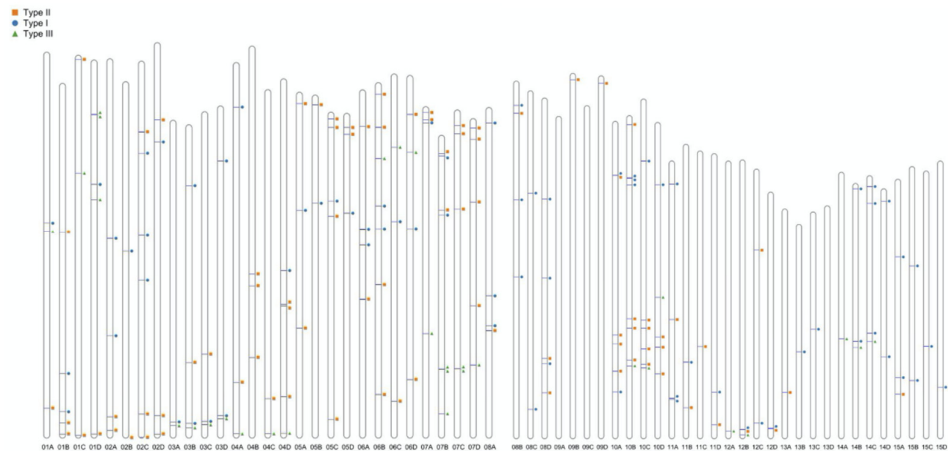


Figure 1. The chromosomal distribution of WRKY genes in the tetraploid *C. oleifera* genome. Each label represents a WRKY gene and the color and shape indicate their types. Of the 192 identified WRKY genes, four genes located on scaffolds were excluded in the visualization.

Table 1. The chromosomal distribution of WRKY in tetraploid *Camellina Oleifera*.

Homolog Group	Subgenome				T _{group}
	A	B	C	D	
1	3	5	3	5	16
2	4	2	6	4	16
3	2	4	3	3	12
4	3	3	2	5	13
5	3	2	5	3	13
6	4	7	3	4	18
7	4	7	5	5	21
8	4	4	2	5	15
9	0	1	0	1	2
10	6	8	6	5	25
11	4	2	1	2	9
12	1	3	2	2	8
13	1	1	1	0	3
14	1	3	4	2	10
15	3	2	1	1	7
T _{subgenome}	43	54	44	47	188

3.2. Conserved Domain and Gene Structure Analysis

In general, WRKY proteins were characterized for two typical structures: WRKY domain and zinc-finger configuration. Based on their structural characteristics, these genes were classified into three major groups (Type I, II and III). Of them, Type I (74 genes, 39% of total) contains single core WRKY domain without zinc-finger motifs; Type II (87 genes, 45% of total) are featured with one or two WRKY domains coupled with C2H2-type zinc-fingers; Type III (31 genes, 16% of total) possesses single WRKY domains and distinct C2HC-type zinc-fingers (Figure 2A). The gene structure analysis showed that the size of WRKY genes varied significantly (1,281 - 23,094 bp) with an average length of around 4,728 bp, and no WRKY gene is intron-less (Figure 2B). Furthermore, no new distinct pattern was found for gene structure among the types (Figure 2B).

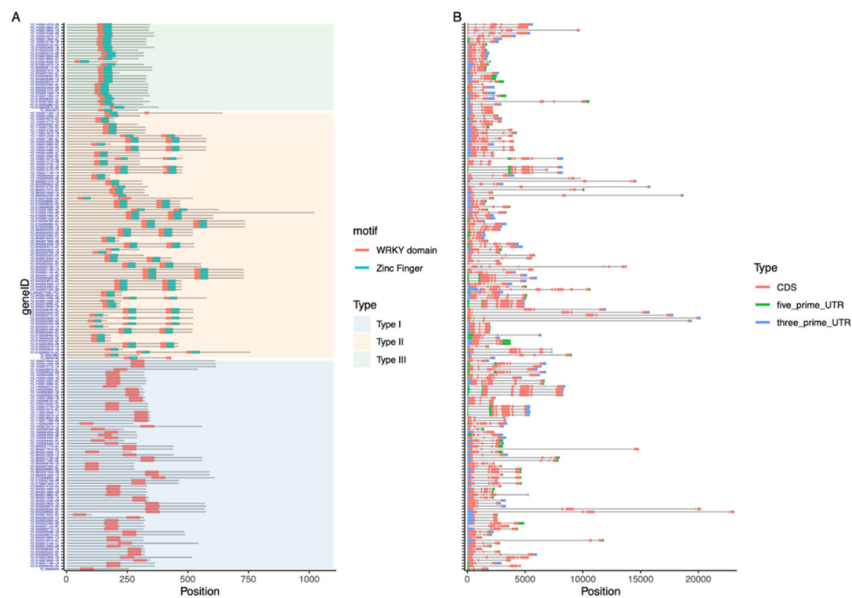


Figure 2. Conserved domains and gene structure of the 192 WRKY genes in tetraploid *C. oleifera* genome. (A) The WRKY was grouped into three types (I to III), highlighted from bottom to top. The colorful boxes indicate WRKY domain and zinc-finger motif; (B) The colorful boxes represent coding sequences (CDS) and untranslated regions (UTR), and gray lines indicate introns.

3.3. Phylogenetic Relationship of the WRKY Genes

A phylogenetic tree was constructed with the full coding proteins to illustrate the relationship among WRKY genes, which have been categorized into three types (I, II, III) based on domain analysis. The tree was composed of three major clusters, where the genes belonging to different types were grouped into, and the genes belonging to the same type were grouped into one sub-cluster (Figure 3). In specific, Type I genes (blue) exhibit significant sub-clustering, indicating recent diversification. Type II genes (orange) show some overlap with Type I, probably suggesting ancient duplication events. Type III genes (green) are more divergent, hinting at unique functional adaptations (Figure 3).

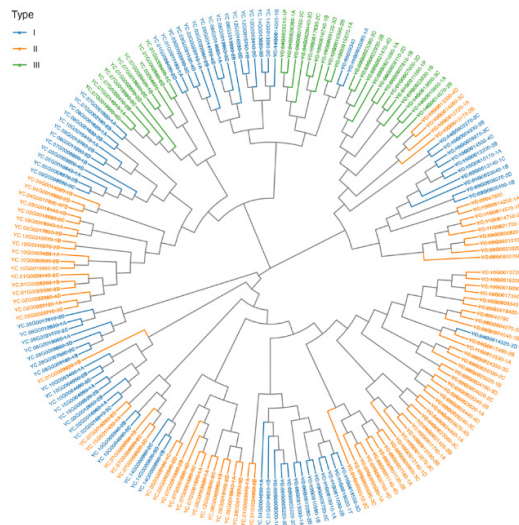


Figure 3. Phylogenetic relationship of WRKY genes in *C. oleifera*. Blue: Type I WRKY; Orange: Type II WRKY; Green: Type III WRKY.

3.4. Gene Expression Pattern of WRKY and Response to Anthracnose Infection

To examine the expression pattern of WRKY genes in response to biotic stress, two *C. oleifera* cultivars with contrasting disease resistance, anthracnose-resistant CL150 and susceptible CL102, were selected for expression profiling under control and pathogen stress. After obviously low expressed WRKY genes were removed, a total of 156 genes were investigated. In general, 156 WRKY genes showed different expression pattern and can be classified into six clusters. Two clusters (I and IV, in Figure 4) displayed moderate and stable expression levels across two cultivars under two conditions, indicating a constitute expression pattern. Cluster II showed relatively higher expression levels in both cultivars and up-regulated under pathogen challenge, suggesting a potential role in defense mechanisms (Figure 4). Two clusters (III and VI) exhibited no expression across two cultivars under control but slightly up-regulated in response to pathogen attack, which could contribute to disease defense. Additionally, the genes in Cluster V performed constant but low expression levels across two cultivars and conditions (Figure 4).

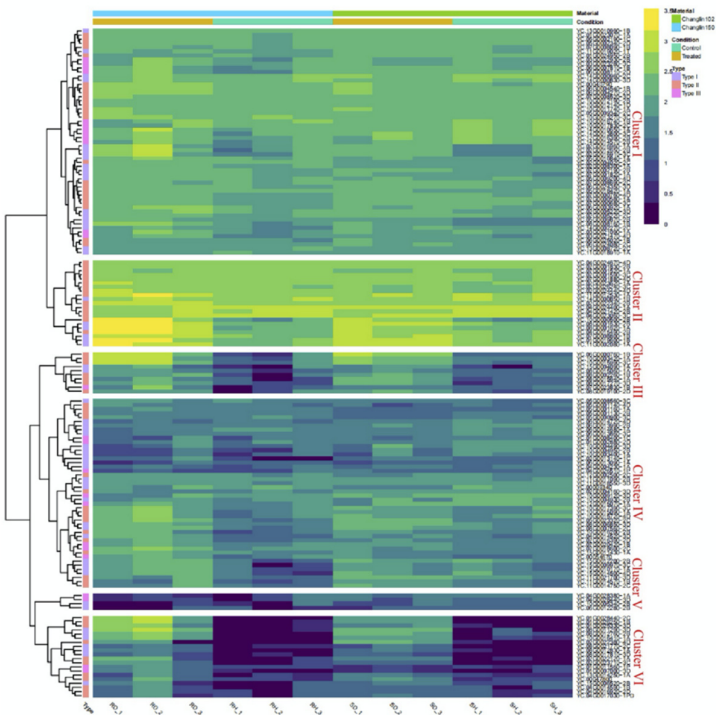


Figure 4. Expression pattern of WRKY gene family. For each sample (column), the expression was profiled for 152 genes (lowly expressed WRKY genes excluded). All genes were classified into 6 clusters (I to VI from top to bottom) based on the expression pattern. RD = Resistant line under disease treatment; RH = Resistant line without treatment; SD = Susceptible line under disease treatment; SH = Susceptible line without treatment.

3.5. Differentially Expressed Genes (DEGs) Under Anthracnose Stress

To explore potential WRKY genes associated with resistance to anthracnose, different expression analysis was conducted on two cultivars under both control and pathogen treatment. A total of 12 samples including biological replicates were investigated for expression level. First, the principal component analysis (PCA) on all samples demonstrated that both genotypes/cultivars and treatment effect contribute to the variation of gene expression levels (Supplementary Figure S1). Second, a total of 1,822 differentially expressed genes (DEGs) were detected in response to anthracnose by both cultivars (Figure 5). Functional enrichment analysis revealed that these genes mainly function as enzymes such as hydroxylase and oxidase (Supplementary Figure S2). Furthermore, a large proportion of them were found to be involved into the defense related biological processes such as cysteine biosynthetic process from serine

(GO:0006535), glutathione metabolic process (GO:0006749), hydrogen peroxide catabolic process (GO:0042744) as well as defense response to bacterium (GO:0042742) (Supplementary Figure S2).

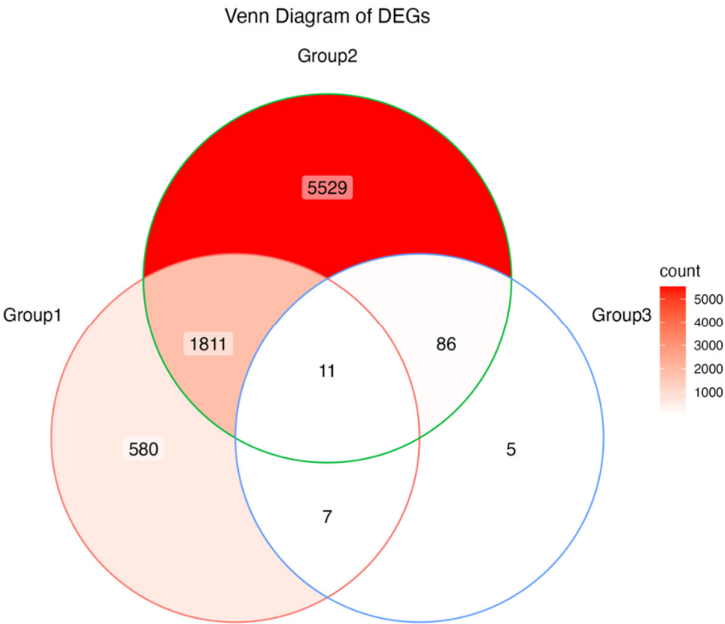


Figure 5. The number of differentially expressed genes (DEGs) across two cultivars between control and disease treatment. Group1 and Group2 represent the DEGs between treatment and control in CL102 and CL150 respectively, and the overlap of these two group (n=1,822) indicate the DEGs in both cultivars under disease treatment; Group3 represents the DEGs (n=109) differed in two cultivars under disease treatment.

3.6. A Candidate WRKY Gene Involved into Resistance to Anthracnose

In order to investigate DEGs distinct to the resistant material (CL150), a further analysis was performed to compare CL102 and CL150 subjected to the disease, the number of 109 DEGs were detected and showed distinct expression pattern (Figure 5, Supplementary Figure S3). These genes mainly function as phosphatase, oxygenase and transcriptional factor involved in the biological processes including inositol catabolism and protein dephosphorylation (Figure 6A). Furthermore; KEGG analysis revealed most of these genes are responsible for the pathways such as hydroperoxide dehydration, glutamine-hydrolysing and dioxygenation of glucoside (Figure 6B). Regarding to expression regulation, 83 genes of 109 DEGs (76%) were up-regulated by 1.2-9.1 folds with the gene YC.14G0015890-3C showing the highest increase in expression level (Figure 7A). Conversely, 26 genes were down-regulated by 1.3-13.2 folds with the gene YC.12G0009430-1A showing the most significant decrease in expression level (Figure 7A).

Of the 109 DEGs, 11 core DEGs (ten of them up-regulated) were observed, representing a group of cultivar-independent DEGs because they were significantly regulated in both cultivars under the disease stress (Figure 5). Furthermore, the function enrichment showed that four of them are associated with reduction of oxygen (Table 2). Interestingly, one WRKY gene (YC.08G0001620-1A, Type I) among those 11 core genes exhibited a significant up-regulation in CL150 (fold-change = 2.2, padj=9.1e-03) compared to that of CL102, although it was also up-regulated in response to the disease stress (Figure 7B).

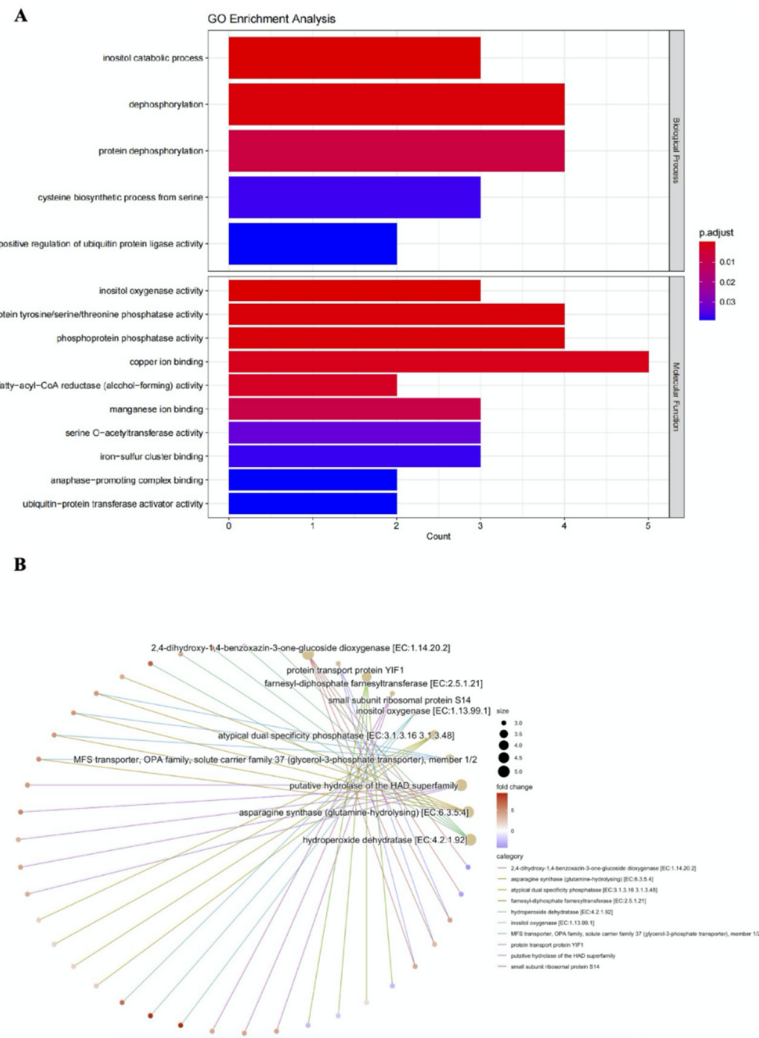


Figure 6. Functional annotation (A) and network analysis (B) on 109 DEGs differed in CL102 and CL150 responded to disease.

Table 2. The expression change and functional annotation on the 11 core DEGs.

Gene ID	Fold Change	p.adj Value	Function Description
YC.01G0018630-2C	6.60	7.03E-03	manganese ion binding
YC.02G0003640-2D	4.38	5.20E-03	acyltransferase activity, transferring groups other than amino-acyl groups
YC.02G0003770-1C	3.95	1.39E-04	acyltransferase activity, transferring groups other than amino-acyl groups
YC.11G0011070-2C	8.56	2.19E-05	alcohol-forming very long-chain fatty acyl-CoA reductase activity
YC.11G0010960-1B	7.74	4.82E-07	alcohol-forming very long-chain fatty acyl-CoA reductase activity
YC.05G0019060-3C	6.54	5.20E-03	unknown
YC.08G0001620-1A	2.23	9.10E-03	DNA-binding transcription factor activity
YC.09G0021520-3C	7.07	1.89E-04	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen
YC.09G0020800-2B	4.37	4.23E-03	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen

YC.09G0022420-1A	5.75	5.80E-03	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen
YC.15G0010630-1C	-2.87	1.69E-03	oxidoreductase activity

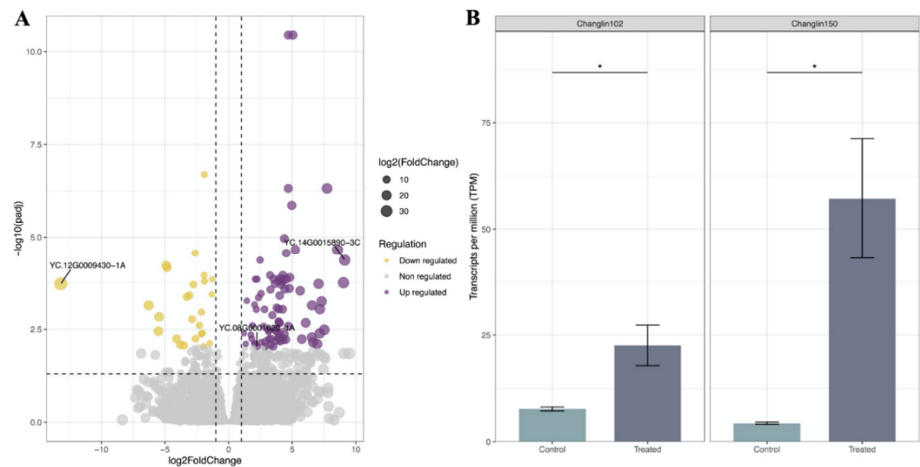


Figure 7. Different expression analysis of *C. oleifera* under two conditions. (A) Volcano plot of 109 DEGs. The 109 DEGs are differentially expressed genes specific to resistant cultivar CL150 (Changlin150) in case of disease treatment. The most up-regulated (right), down-regulated (left) and WRKY gene (middle) were labeled with their gene names. (B) YC.08G0001620-1A expression response to disease treatment. The expression levels are represented by the read counts, and student t-test was used for statistical analysis (*, significant level 0.05).

4. Discussion

4.1. Expansion and Evolutionary Dynamics of the WRKY Family

The number of WRKY genes in *C. oleifera* (n=192) exceeds that of diploid species like *Arabidopsis thaliana* (n=72) and *Oryza sativa* (n=101) [34], suggesting lineage-specific expansion, likely driven by whole-genome duplication (WGD) events common in polyploid plants [35]. Although the WGD led to a wide distribution for WRKY genes in the tetraploid *C. oleifera* genome, hotspots are still present for gene localization like other polyploid species like soybean and cotton [21, 22]. For instance, Chr10 (-A, -B, -C and -D) harbor the highest density of WRKY genes (n=25), indicating a consequence of segmental duplication. Conversely, the absence of WRKY genes on Chr09A, Chr09C, and Chr13D may reflect chromosomal rearrangements post-polyploidization. Although the number of WRKY genes carried on each chromosome varies considerably (n=3 to 25), the number of WRKY genes in each subgenome shows little variation (n=44 to 54). This phenomenon reflects dosage-balance constraints that maintain a constant functional reservoir of this key transcription factor family during polyploidization, as shown in *Liriodendron chinense* [36].

The mixed clusters of various WRKY types (e.g. Type II genes overlapping with Type I), showed in the phylogenetic tree, suggests ancient duplication events followed by subfunctionalization. Similar patterns in *Oryza sativa* [37] and *Populus trichocarpa* [38] have been linked to adaptive evolution under biotic stress. The distinct divergence of Type III genes, forming a separate clade, may reflect unique selective pressures, potentially linked to co-evolution with pathogen [39].

4.2. Expression Dynamics on Anthracnose Resistance

The expression pattern of WRKY genes in resistant (CL150) and susceptible (CL102) cultivars underscores their roles in disease response. The genes in Cluster III and Cluster VI, up-regulated under pathogen challenge, likely participate in defense signaling, like WRKY18/40/60 in Arabidopsis, which activate jasmonate/ethylene pathways [40]. The highly induced expression of Cluster II genes under stress suggests roles in pathogen recognition or early defense. Notably, the candidate gene *YC.08G0001620-1A* (Type I), showed significant up-regulation in both resistant and susceptible cultivars. The conserved up-regulation of *YC.08G0001620-1A* in *C. oleifera* implies a conserved role in priming defense responses, potentially through transcriptional activation of pathogenesis-related (PR) genes.

4.3. Candidate Genes and Process to Anthracnose Resistance

The 109 cultivar-specific DEGs, particularly the 11 core DEGs, highlight core regulatory hubs in anthracnose resistance. The strong enrichment into activities like DNA-binding and oxidoreductase, implying that *YC.08G0001620-1A* may be potentially as a master regulator to coordinate reductase of reactive oxygen species (ROS), a similar mechanism in rapeseed where expression of one WRKY gene *BnaWGR1* induced an accumulation of ROS both in tobacco and Arabidopsis [41]. To validate its role, further studies are required for functional characterization via CRISPR knockout or overexpression in model systems, as demonstrated for *Vitis vinifera* in powdery mildew resistance [42].

An extensive study was conducted for identification of the WRKY transcription factor family in the tetraploid *C. oleifera* genome and further identified a key WRKY candidate (*YC.08G0001620-1A*), as a regulatory factor associated with anthracnose resistance. However, critical questions remain unresolved: (1) whether this WRKY gene confers broad-spectrum resistance against multiple pathogens or is specific to *Colletotrichum gloeosporioides*, and (2) which downstream functional genes or regulatory modules interact with this transcription factor to orchestrate defense responses. Addressing these questions will require further investigations employing such as multi-omics approaches (e.g., CHIP-Seq and co-expression network analysis) combined with functional validation in model systems.

5. Conclusions

This study provides a comprehensive overview of the WRKY gene family in *C. oleifera* and identifies a potential candidate gene (*YC.08G0001620-1A*) for anthracnose resistance. Future studies should focus on its functional characterization through genetic transformation and biochemical assays to elucidate its role in disease resistance. Additionally, further investigation of other differentially expressed genes may reveal additional mechanisms underlying anthracnose resistance in *C. oleifera*. This research contributes to the understanding of the molecular basis of disease resistance in *C. oleifera* and provides valuable information for breeding programs aimed at improving anthracnose resistance in this important oil crop.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: title; Table S1: title; Video S1: title.

Author Contributions: Conceptualization, J.C. and J.P.; methodology, J.C. and K.O.; formal analysis, D.J. and Y.X.; investigation, J.C.; resources, Y.D.; data curation, Y.D. and Y.J.; writing—original draft preparation, J.C. and K.O.; writing—review and editing, J.C. and J.P.; visualization, Y.X.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work is financially supported by Hubei Natural Science Foundation (2025AFB961), Key Research and Development Program of Hubei Province (2025BBB010) and Foundation of Hubei Academy of Forestry (2025-ZZLX-A08).

Data Availability Statement: The original data presented in the study are openly available in the NCBI repository (accession: PRJNA775660).

Acknowledgments: We gratefully thank Yongqing Cao at Research Institute of Subtropical Forestry (RISF), Chinese Academy of Forestry (CAF) for providing plant materials.

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

References

1. Song, H., et al., Review: WRKY transcription factors: Understanding the functional divergence. *Plant Sci*, 2023. 334: p. 111770.
2. Eulgem, T., et al., The WRKY superfamily of plant transcription factors. *Trends Plant Sci*, 2000. 5(5): p. 199-206.
3. Rushton, P.J., et al., WRKY transcription factors. *Trends Plant Sci*, 2010. 15(5): p. 247-58.
4. Yin, W., et al., Overexpression of VqWRKY31 enhances powdery mildew resistance in grapevine by promoting salicylic acid signaling and specific metabolite synthesis. *Hortic Res*, 2022. 9.
5. Liu, Y., et al., MdERF114 enhances the resistance of apple roots to *Fusarium solani* by regulating the transcription of MdPRX63. *Plant Physiol*, 2023. 192(3): p. 2015-2029.
6. Wang, Y., et al., VqWRKY56 interacts with VqbZIPC22 in grapevine to promote proanthocyanidin biosynthesis and increase resistance to powdery mildew. *New Phytol*, 2023. 237(5): p. 1856-1875.
7. Mu, Y., et al., JrPHL8-JrWRKY4-JrSTH2L module regulates resistance to *Colletotrichum gloeosporioides* in walnut. *Hortic Res*, 2024. 11(7): p. uhae148.
8. Yan, H., et al., The cotton WRKY transcription factor GhWRKY17 functions in drought and salt stress in transgenic *Nicotiana benthamiana* through ABA signaling and the modulation of reactive oxygen species production. *Plant Cell Physiol*, 2014. 55(12): p. 2060-76.
9. Wei, W., et al., GmWRKY54 improves drought tolerance through activating genes in abscisic acid and Ca(2+) signaling pathways in transgenic soybean. *Plant J*, 2019. 100(2): p. 384-398.
10. Zhang, J., et al., Drought-responsive WRKY transcription factor genes IgWRKY50 and IgWRKY32 from *Iris germanica* enhance drought resistance in transgenic *Arabidopsis*. *Front Plant Sci*, 2022. 13: p. 983600.
11. Yuan, S., et al., BcNAC056 Interacts with BcWRKY1 to Regulate Leaf Senescence in Pak Choi. *Plant Cell Physiol*, 2023. 64(9): p. 1091-1105.
12. Zhao, L., et al., A WRKY transcription factor, TaWRKY40-D, promotes leaf senescence associated with jasmonic acid and abscisic acid pathways in wheat. *Plant Biol (Stuttg)*, 2020. 22(6): p. 1072-1085.
13. Singh, D., et al., Tomato (*Solanum lycopersicum*) WRKY23 enhances salt and osmotic stress tolerance by modulating the ethylene and auxin pathways in transgenic *Arabidopsis*. *Plant Physiol Biochem*, 2023. 195: p. 330-340.
14. Chen, J., et al., *Arabidopsis* VQ10 interacts with WRKY8 to modulate basal defense against *Botrytis cinerea*. *J Integr Plant Biol*, 2018. 60(10): p. 956-969.
15. Huang, H., et al., SIVQ15 interacts with jasmonate-ZIM domain proteins and SIWRKY31 to regulate defense response in tomato. *Plant Physiol*, 2022. 190(1): p. 828-842.
16. Yang, C., et al., Comparison of Oil Content and Fatty Acid Profile of Ten New *Camellia oleifera* Cultivars. *J Lipids*, 2016. 2016: p. 3982486.
17. Wang, Y., et al., Identification and Characterization of *Colletotrichum* Species Associated with Anthracnose Disease of *Camellia oleifera* in China. *Plant Dis*, 2020. 104(2): p. 474-482.
18. Zhang, L., et al., The tetraploid *Camellia oleifera* genome provides insights into evolution, agronomic traits, and genetic architecture of oil *Camellia* plants. *Cell Rep*, 2024. 43(12): p. 115032.
19. Zhu, H., et al., The complex hexaploid oil-*Camellia* genome traces back its phylogenomic history and multi-omics analysis of *Camellia* oil biosynthesis. *Plant Biotechnol J*, 2024. 22(10): p. 2890-2906.
20. Wang, R., et al., Haplotype-resolved genome assembly of the tetraploid Yousha tree *Camellia meiocarpa* Hu. *Sci Data*, 2025. 12(1): p. 541.
21. Dou, L., et al., Genome-wide analysis of the WRKY gene family in cotton. *Mol Genet Genomics*, 2014. 289(6): p. 1103-21.
22. Yu, Y., et al., Genome-wide identification of soybean WRKY transcription factors in response to salt stress. *Springerplus*, 2016. 5(1): p. 920.

23. Yang, C., et al., Integrated Transcriptome and Metabolome Analysis Reveals Key Metabolites Involved in *Camellia oleifera* Defense against Anthracnose. *Int J Mol Sci*, 2022. 23(1).
24. Eddy, S.R., Accelerated Profile HMM Searches. *PLoS Comput Biol*, 2011. 7(10): p. e1002195.
25. Jones, P., et al., InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 2014. 30(9): p. 1236-40.
26. Hao, Z., et al., RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the ideograms. *PeerJ Comput Sci*, 2020. 6: p. e251.
27. Thompson, J.D., D.G. Higgins, and T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, 1994. 22(22): p. 4673-80.
28. Stecher, G., K. Tamura, and S. Kumar, Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol Biol Evol*, 2020. 37(4): p. 1237-1239.
29. Chen, S., et al., fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 2018. 34(17): p. i884-i890.
30. Li, H. and R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009. 25(14): p. 1754-60.
31. Liao, Y., G.K. Smyth, and W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 2014. 30(7): p. 923-30.
32. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 2014. 15(12): p. 550.
33. Wu, T., et al., clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*, 2021. 2(3): p. 100141.
34. Abdullah-Zawawi, M.R., et al., Comparative genome-wide analysis of WRKY, MADS-box and MYB transcription factor families in *Arabidopsis* and rice. *Sci Rep*, 2021. 11(1): p. 19678.
35. Du, C., et al., Evolutional, expressional and functional analysis of WRKY gene family reveals that PbeWRKY16 and PbeWRKY31 contribute to the Valsa canker resistance in *Pyrus betulifolia*. *Plant Physiol Biochem*, 2025. 222: p. 109719.
36. Wu, W., et al., Genome-wide identification of the *Liriodendron chinense* WRKY gene family and its diverse roles in response to multiple abiotic stress. *BMC Plant Biology*, 2022. 22(1): p. 25.
37. Ryu, H.S., et al., A comprehensive expression analysis of the WRKY gene superfamily in rice plants during defense response. *Plant Cell Rep*, 2006. 25(8): p. 836-47.
38. Chen, Y., et al., Genome-Wide Identification of WRKY Family Genes and the Expression Profiles in Response to Nitrogen Deficiency in Poplar. *Genes (Basel)*, 2022. 13(12).
39. Yin, G., et al., The large soybean (*Glycine max*) WRKY TF family expanded by segmental duplication events and subsequent divergent selection among subgroups. *BMC Plant Biol*, 2013. 13: p. 148.
40. Chen, H., et al., Roles of *Arabidopsis* WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biol*, 2010. 10: p. 281.
41. Yang, L., et al., An oilseed rape WRKY-type transcription factor regulates ROS accumulation and leaf senescence in *Nicotiana benthamiana* and *Arabidopsis* through modulating transcription of RbohD and RbohF. *Planta*, 2018. 247(6): p. 1323-1338.
42. Zhou, M., et al., Transcription factors VviWRKY10 and VviWRKY30 co-regulate powdery mildew resistance in grapevine. *Plant Physiol*, 2024. 195(1): p. 446-461.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.