


# Genome-Wide Identification and Characterization of the *WRKY* Gene Family and Their Associated Regulatory Elements in *Fortunella hindsii*

Evolutionary Bioinformatics  
Volume 21: 1–13  
© The Author(s) 2025  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/11769343241312740



Hadia Hussain<sup>1,2</sup> , Aleena Alam<sup>3</sup>, Iqra Mehar<sup>4</sup>, Maryam Noor<sup>2</sup>, Othman Al-Dossary<sup>5</sup>, Bader Alsubaie<sup>5</sup>, Muneera Q. Al-Mssallem<sup>6</sup> and Jameel Mohammed Al-Khayri<sup>5</sup>

<sup>1</sup>Key Laboratory of Molecular Epigenetics of MOE, Northeast Normal University, Changchun, China. <sup>2</sup>Department of Biotechnology, University of Okara, Okara, Pakistan. <sup>3</sup>Department of Agricultural Entomology and Pest Control, College of Plant Protection, Jilin Agricultural University, Jilin, Changchun, China. <sup>4</sup>Government College University Faisalabad, Faisalabad, Pakistan. <sup>5</sup>Department of Agricultural Biotechnology, College of Agriculture and Food Sciences, King Faisal University, Al-Ahsa, Saudi Arabia. <sup>6</sup>Department of Food Science and Nutrition, College of Agriculture and Food Sciences, King Faisal University, Al-Ahsa, Saudi Arabia.

## ABSTRACT

**BACKGROUND:** The WRKY gene family is identified as one of the most prominent transcription factor families in plants and is involved in various biological processes such as metabolism, growth and development, and response to biotic and abiotic stresses. In many plant species, the WRKY gene family was widely studied and analyzed but little to no information for *Fortunella hindsii*. However, the completion of the whole genome sequencing of *Fortunella hindsii* allowed us to investigate the genome-wide analysis of WRKY proteins.

**OBJECTIVE:** The main objective of this study was to analyze and identify the WRKY gene family in *Fortunella hindsii* genome.

**METHODOLOGY:** Various bioinformatics approaches have been used to conduct this study.

**RESULTS:** We constituted 46 members of the *Fortunella hindsii* WRKY gene family, which were unevenly distributed on all nine chromosomes. The phylogenetic relationship of predicted WRKY proteins of *Fortunella hindsii* with the WRKY proteins of *Arabidopsis* showed that 46 *FhWRKY* genes were divided into three main groups (G1, G2, G3) with five subgroups (2A, 2B, 2C, 2D, and 2E) of G2 group. Domain, conserved motif identification, and gene structure were conducted and the results found that these *FhWRKY* proteins have conserved identical characteristics within groups and maintain differences between groups. In silico subcellular localization, results showed that *FhWRKY* genes are located in the nucleus. The *cis*-regulatory element analysis identified several key CREs that are significantly associated with light, hormone responses, and stress. The gene ontology analysis of these predicted *FhWRKY* genes showed that these genes are significantly enriched in sequence-specific DNA binding, transcriptional activity, cellular biosynthesis, and metabolic processes.

**CONCLUSION:** Therefore, overall, our results provided an excellent foundation for further functional characterization of WRKY genes with an aim of *Fortunella hindsii* citrus crop improvement.

**KEYWORDS:** *Fortunella hindsii*, WRKY transcription factor, genome-wide analysis, bioinformatics analysis, abiotic stresses

**RECEIVED:** August 8, 2024. **ACCEPTED:** December 16, 2024.

**TYPE:** Original Research

**FUNDING:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KF250922].

**DECLARATION OF CONFLICTING INTERESTS:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**CORRESPONDING AUTHORS:** Hadia Hussain, Key Laboratory of Molecular Epigenetics of MOE, Northeast Normal University, Changchun 130024, China. Email: hadiahussain37@outlook.com

Jameel Mohammed Al-Khayri, Department of Agricultural Biotechnology, College of Agriculture and Food Sciences, King Faisal University, Al-Ahsa 31982, Saudi Arabia. Email: jkhayri@kfu.edu.sa

## Introduction

As a sessile organism, plants are exposed to many environmental factors, including biotic and abiotic stresses such as heat, cold, drought, and high salinity, and have diverse mechanisms to alleviate the effects of these environmental fluctuations.<sup>1</sup> Transcription factors (TFs) play a significant role in response to dynamic extrinsic and intrinsic stimuli to regulate gene expression.<sup>2</sup>

The WRKY family was found to be one of the most significant families of transcription factors in higher plants.<sup>3</sup> The WRKY term is derived from four highly conserved amino

acids in its domain. Structurally, the WRKY proteins have one or two DNA binding domains consisting of approximately 60 amino acids with a conserved heptapeptide WRKYGQK stretch at its N-terminal and a zinc-finger motif (CX4-5CX22-23HXX or CX7Cx23HXC) at its C-terminal.<sup>4</sup> WRKY proteins are classified into three distinct groups (Group I, Group II, and Group III) based on the specific type of their zinc-finger motifs and several conserved domains. Moreover, Group II can be further classified into five subgroups (Group II a-e) based on their evolutionary clades and distinct assemblies.<sup>5</sup>



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

The genome-wide identification of WRKY TFs has been reported in various plant species including the model plant *Arabidopsis thaliana* (72–74), *Zea mays* (136), *Oryza sativa* (102), *Glycine max* (182), *Citrus sinensis* (51), and *Citrus clementine* (48).<sup>6–8</sup> WRKY proteins have many biological activities due to their unique domains. In 1994, the WRKY transcription factor SPF1 was first characterized from sweet potato.<sup>9</sup> Previous studies have reported that WRKY transcription factors are involved in seed size,<sup>10</sup> hair development,<sup>11</sup> flowering time,<sup>12</sup> pollen development,<sup>13</sup> leaf senescence,<sup>14</sup> and plant hormone signaling pathways.<sup>15,16</sup> For instance, *AtWRKY46* in *Arabidopsis* regulates the lateral root development during osmotic/salt stress conditions through the regulation of auxin homeostasis and ABA signaling.<sup>3</sup> In sugarcane (*Saccharum* spp.), transcription factor WRKY class III was involved in response to abiotic and biotic stresses.<sup>17</sup>

Despite their diverse roles in plant signaling pathways, growth and development, the most significant function of WRKY proteins is transcriptional regulation in response to biotic and abiotic stresses.<sup>18</sup> Numerous WRKY transcription factors had their expressions controlled by *Arabidopsis* *AGB1*, and it is responsible for mediating the primary metabolism and stress-responsive genes under excess nitrogen. WRKYs have distinct roles, as shown by gene ontology and mutant analysis. While *AtWRKY75* controls metabolic activities, *AtWRKY40* is involved in ABA and typical stress responses.<sup>19</sup>

Worldwide, citrus fruit is one of the most widely cultivated fruit crops. The Hongkong kumquat (*Fortunella hindsii*) is a wild species of citrus classified within the Citrinae group in the subfamily Aurantioideae of the family Rutaceae.<sup>20</sup> This wild fruit tree is native to southern China. It is distinguished from other Citrus trees by its smaller size, which ranges from 20 to 120 cm, and its earlier flowering period of approximately 8 months. Moreover, it might bloom twice or three times a year.<sup>21</sup> *Fortunella hindsii* have significant agronomic characteristics such as compact tree form, an edible peel, and cold resistance.<sup>22</sup> Recent studies have shown that the yield of citrus fruit crops is gradually decreasing due to the increasing effects of global warming, including biotic and abiotic stresses such as cold, drought, diseases, and salinity.<sup>23</sup> Over the years traditional methods have been used for citrus crop improvements. However, these approaches are limited because of incompatibility, parthenocarp, stunted growth, and polyembryony.<sup>20</sup> To overcome these limitations different biological approaches and computational tools are used. Several WRKY genes were reported to regulate response to pathogens and abiotic stresses in many species of citrus such as *C. sinensis*, *C. unshiu*, *C. reticulata*, and *C. clementina*.<sup>8</sup> The citrus WRKY family may have a potential role in a variety of development processes in plants and responses to biotic and abiotic stresses. Therefore, the WRKY family members could be potential candidates for improvement programs and citrus breeding.

In this study, various bioinformatics tools were used to analyze the WRKY gene family comprehensively in *Fortunella*

*hindsii*. To identify members of WRKY genes in the genome, a deterministic strategy was used including intron/exon distribution, conserved motifs, chromosome distribution, comparative phylogenetic analysis, sequence alignment, structure prediction, cis-regulatory elements, and enrichment analysis. Moreover, *Fortunella hindsii* genome-wide analysis of the WRKY gene family served as a vulnerable reference and suitability for practical analysis and the cloning of specifically targeted genes.

## Materials and Methods

### Sequence retrieval of WRKY gene family in the *Fortunella hindsii*

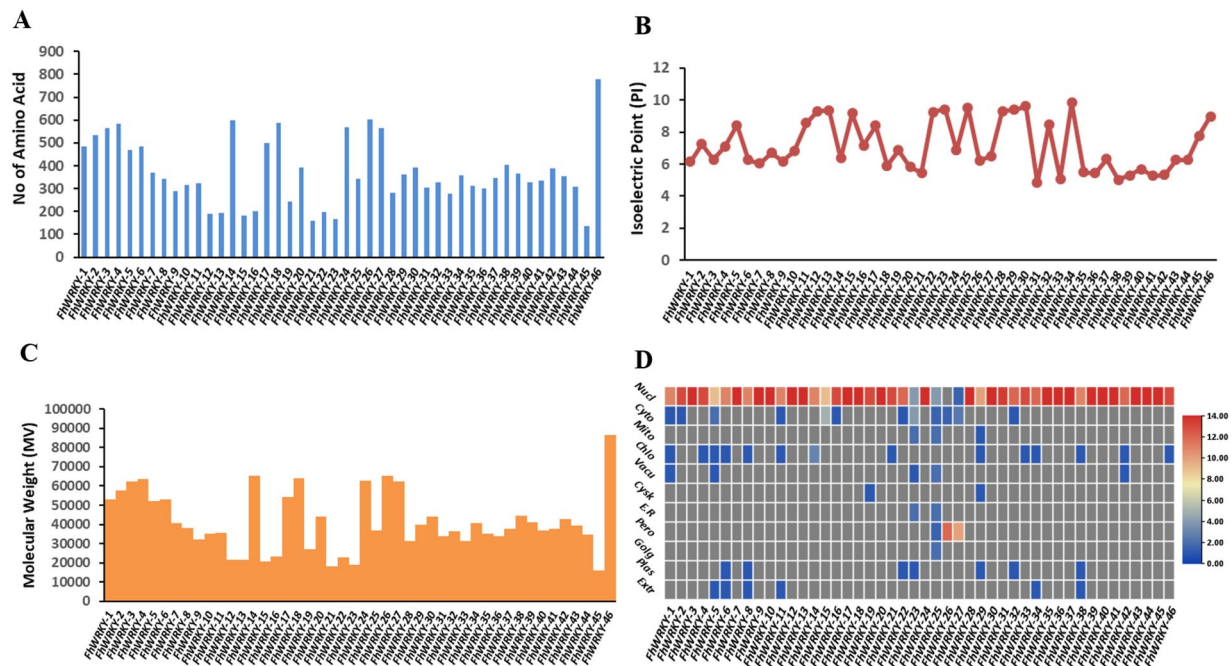
The *Fortunella hindsii* genome files v2.0 were retrieved from the citrus genome database (CGD) (<https://www.citrusgenomedb.org/> accessed on February 3, 2024).<sup>24</sup> The domain sequences of 73 known member genes of the WRKY transcription factor family of *Arabidopsis thaliana* were retrieved from TAIR (<https://www.arabidopsis.org/browse/genefamily/index.jsp/> accessed on February 3, 2024).<sup>25</sup> Next, protein sequences of *Fortunella hindsii* and *Arabidopsis* were subjected to identify FhWRKY proteins by using protein BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi/> accessed on February 6, 2024), and FhWRKY proteins were successfully matched to corresponding *Arabidopsis* proteins. In addition, all the candidate WRKY proteins were further analyzed for the confirmation of the conserved domain in each WRKY protein through NCBI-Batch CD Search (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>, accessed on February 7, 2024). Finally, the genes that only have conserved WRKY domains were selected for subsequent analysis. The physiochemical properties were calculated to predict the protein length (AA residues), molecular weight (MW), and theoretical isoelectric point (pI) through the online ExPASy program (<https://web.expasy.org/protparam/> accessed on February 8, 2024),<sup>26</sup> while subcellular localization was conducted by WOLF PSORT (<https://wolfsort.hgc.jp/> accessed on February 9, 2024).

### Analyzing gene structure and chromosomal location

To predict the gene structure and chromosomal location of all the candidate FhWRKYs, we utilized TBtools for identification and visualization.<sup>27</sup> Each FhWRKY gene was mapped to *Fortunella hindsii* chromosomes according to the gene number and location, where duplicated gene pairs were linked with a red line.<sup>28,29</sup>

### Multiple sequence alignment and phylogenetic tree construction

To further identify the evolutionary relationships between the WRKY genes in *Fortunella hindsii*. The WRKY family can be divided into distinct groups based on the conserved domain of the WRKY proteins of *Fortunella hindsii*. For this purpose, the ClustalW program was carried out to perform the



**Figure 1.** The physicochemical analysis of the predicted FhWRKY proteins: (A) the size of proteins encoded by FhWRKY genes, (B) variation of isoelectric points in the FhWRKYs, (C) the molecular weight of the FhWRKYs proteins, and (D) in silico sub-cellular localization of the FhWRKYs proteins, the sidebar shows highest count present in each category.

multiple sequence alignment analysis of predicted WRKY amino acid residues by using default parameters. Then, a neighbor-joining (NJ) method based phylogenetic tree was constructed by using MEGA 11 software.<sup>30</sup> The iTOL tool was used to annotate and beautify the phylogenetic tree (<https://itol.embl.de/> accessed on February 11, 2024).

#### Cis-regulatory elements and conserved motifs analysis

To analyze *cis*-regulatory elements (CREs) within the promoter region of the FhWRKY gene family, sequences located 1000 bp upstream of the start codon were obtained from the citrus genome database. The CREs of the FhWRKY genes were identified by the PlantPAN 4.0 database (<http://PlantPAN.itps.ncku.edu.tw/>, accessed on July 15, 2024).<sup>31</sup> The conserved motifs of the identified 46 WRKY genes were predicted using MEME-Suite (<https://meme-suite.org/tools/meme>, accessed on February 14, 2024) tool to understand the structure and peptide sequence of FhWRKY genes with default values and maximum number of motif set or other variables.<sup>32</sup>

#### Analysis of synteny

The Multiple Collinearity Scan toolbox (MCScanX) was used to evaluate the duplicated genes event using the default setup.<sup>33</sup> To identify the synteny relationship between the paralogous *Fortunella hindsii* genes and the orthologous genes of FhWRKY in *Poncirus trifoliata*, *C. clementina*, *C. sinensis*, *Oryza sativa*, and *Arabidopsis*, dual synteny maps were made using TBtools.<sup>34</sup>

#### Functional annotation and 3D structure prediction

To explore the functional annotation of the predicted FhWRKY genes, GO ontology was carried out by using Egglog (<http://egglog5.embl.de/> accessed on February 15, 2024) and visualized by WEGO 2.0 (<https://wego.genomics.cn/> accessed on February 15, 2024). The three-dimensional (3D) protein structure for the identified FhWRKY proteins was predicted by using SWISS MODEL (<https://swissmodel.expasy.org/> accessed on February 18, 2024).

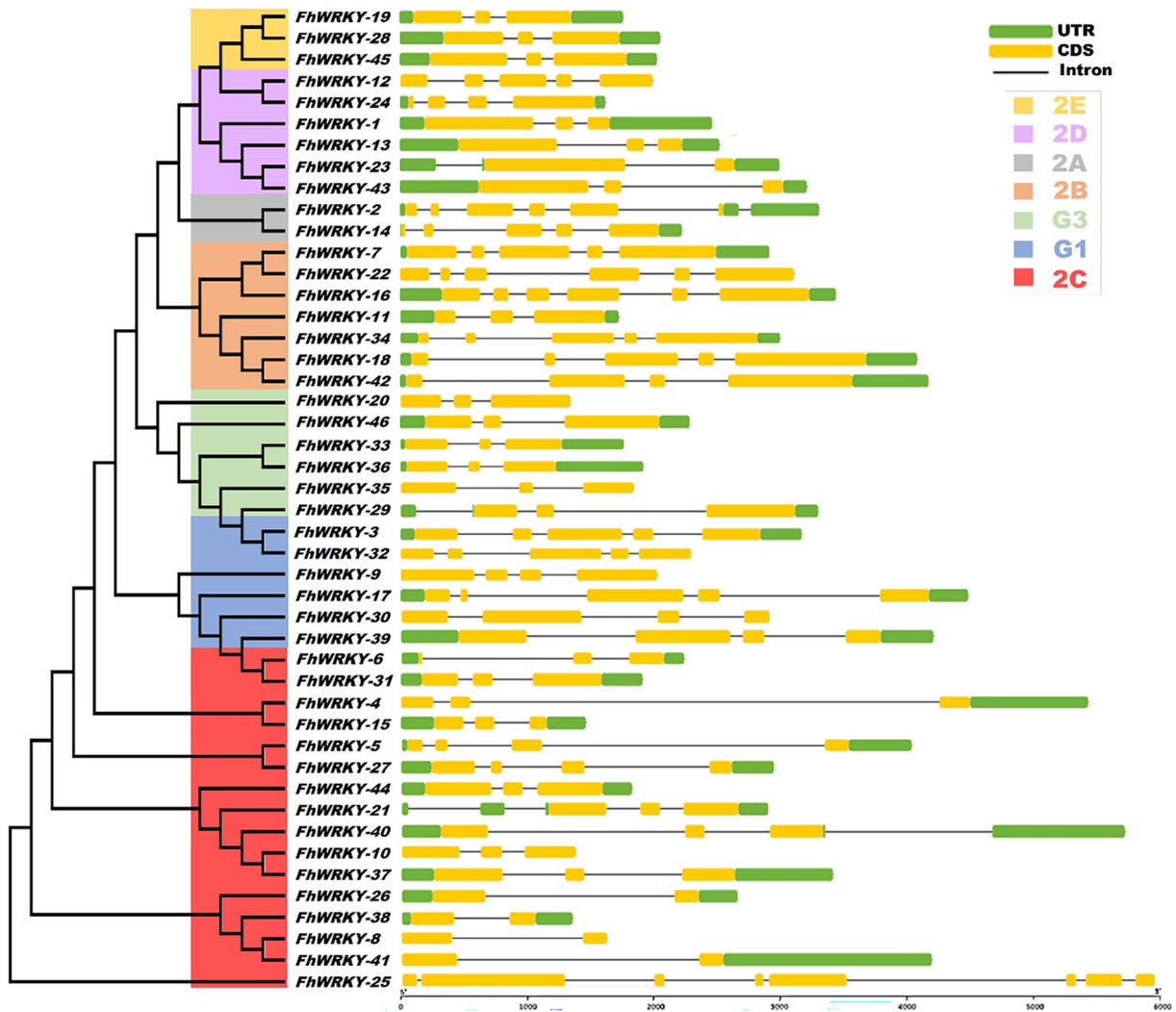
## Results

#### Identification of WRKY TFs in *Fortunella hindsii*

Based on residues of amino acid WRKY transcriptional gene family of *Arabidopsis thaliana*, 46 total FhWRKY genes were successfully obtained from the genome of *Fortunella hindsii* using BLAST-P, after the removal of sequences and duplicates without the corresponding domains. These genes were named as FhWRKY-1 to FhWRKY-46 according to their gene structure and gene ID. Detailed information of the predicted 46 FhWRKY genes is listed in Table S1 (provided in Supplemental File 1). Among all the WRKY proteins in *Fortunella hindsii*, the WRKY domain was 100% highly conserved.

Further, we have analyzed the physicochemical properties of these predicted proteins. In all the predicted FhWRKY proteins, the number of amino acids ranged from 138 to 779 residues (Figure 1A). The molecular weight ranged from 15 903.35 to 86 356.96 kDa (Figure 1C), and the isoelectric point varied from 4.82 to 9.83 (Figure 1B). To further examine where these genes are located in the cell. We have conducted in silico sub-cellular





**Figure 2.** The exon-intron structure of predicted *FhWRKY* genes. Each *FhWRKY* gene length of exon-intron was shown proportionally along with the phylogenetic tree. The number of introns, CDS, and UTR were represented in black lines, yellow bars, and green bar in each gene.

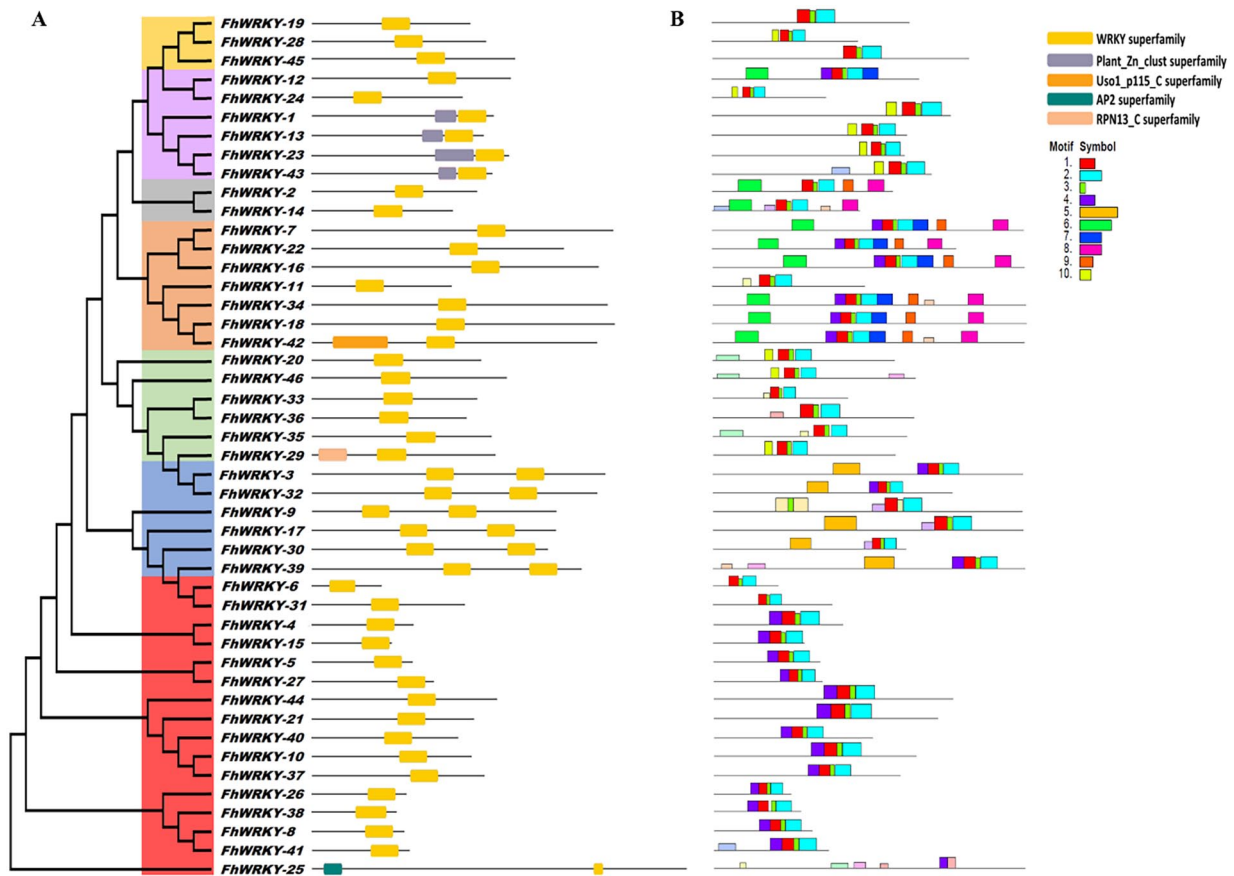
localization analysis by WOLF PSORT.<sup>35</sup> In silico sub-cellular localization results showed that most of the predicted *FhWRKY* genes (96%) were predominantly located in the nucleus except the *FhWRKY23*, *FhWRKY25*, *FhWRKY26*, and *FhWRKY27*, these genes were found in the peroxisome (Figure 1D) detailed data is listed in Table S2 (provided in Supplemental File 1). As previously reported, nuclear localization spotlights the essential involvement of *WRKY* genes in various cellular processes like growth, development, and responses to stress by modulating the expression of target genes.<sup>36</sup>

#### *Analysis of gene structure, conserved functional domains, and motifs in FhWRKYs*

To determine the evolutionary relationships between different genes, the structure and number of exons and introns play a vital role in the identification and characterization process.<sup>37</sup> Therefore, we further investigated the *WRKY* intron-exon coding structure so that the protein expression sequence can be deeply in-sighted. The results showed that

the CDS, UTR, and introns were uniformly distributed in the *FhWRKY* genes (Figure 2). The number of introns in the *FhWRKY* gene family ranged from 1 (*FhWRKY-8*, *FhWRKY-26*, *FhWRKY-38*, *FhWRKY-41*) to 7 (*FhWRKY-25*), with an average number of 2.80. The number of CDSs were consistently distributed throughout the *FhWRKY* genes ranged from 2 (*FhWRKY-8*, *FhWRKY-23*, *FhWRKY-26*, *FhWRKY-38*, *FhWRKY-41*) to 8 (*FhWRKY-25*) and the number of UTRs were found ranging from 1 (*FhWRKY-4*, *FhWRKY-14*, *FhWRKY-41*) to 4 (*FhWRKY-21*), but there is no UTR found in *FhWRKY-8*, *FhWRKY-9*, *FhWRKY-10*, *FhWRKY-12*, *FhWRKY-20*, *FhWRKY-22*, *FhWRKY-25*, *FhWRKY-30*, *FhWRKY-32*, *FhWRKY-35*.

Additionally, a total of 10 identified motifs in *Fortunella hindsi* *WRKY* proteins were analyzed through the MEME program. It was revealed that most of the *FhWRKY* proteins within the same group have similar patterns of motifs and domains, representing that these conserved motifs are involved in the specific group activities among the whole gene family (Figure 3). Motif symbols were represented in different colors



**Figure 3.** The distribution of domains (A) and motifs (B) on 46 *FhWRKY* proteins of *Fortunella hindsii* on the basis of phylogenetic groups to better understand their association. Different color bars represented domains and motifs type in each group of *FhWRKY* genes.

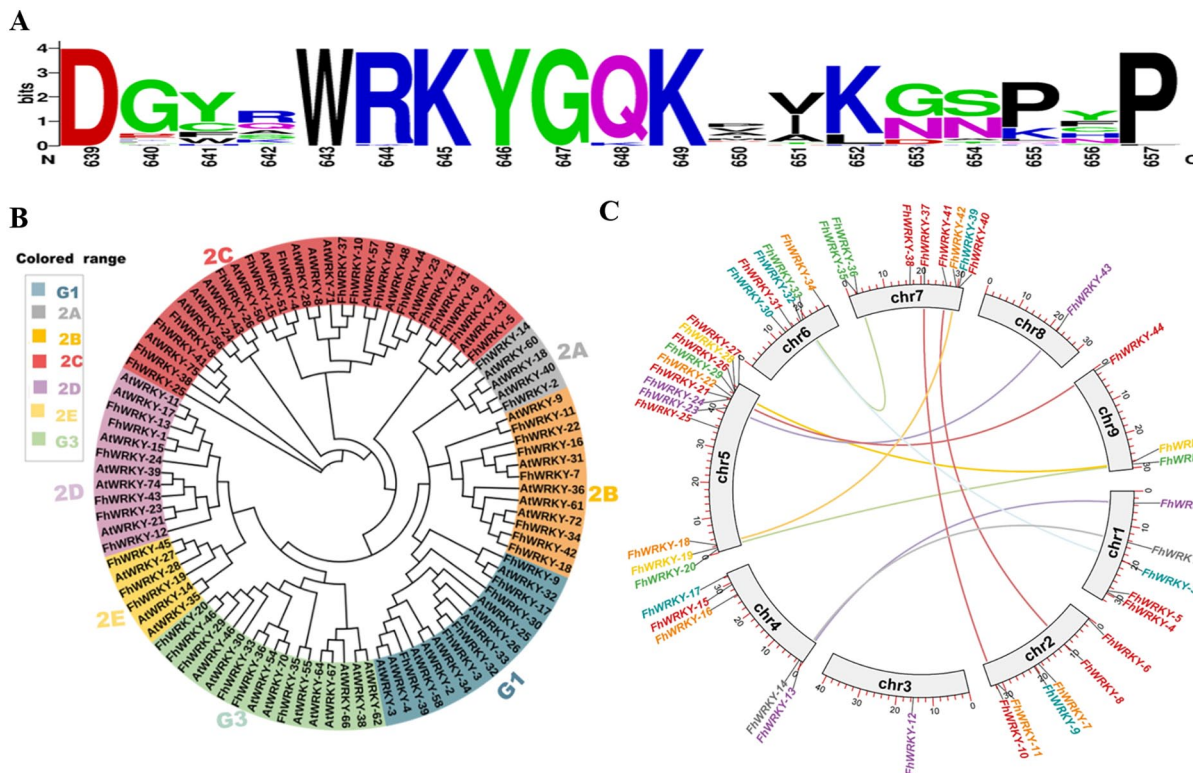
(Figure 3), and it was also identified that motif 1 and motif 5 contain highly conserved WRKYGQK sequence patterns listed in Table S3 (provided in Supplemental File 1).

The domain of WRKY was consistent in all the *FhWRKY* proteins. The two WRKY domains, collectively made up of 5 to 7 motifs present in the group G1 further support the distribution of phylogenetic clade in the *FhWRKY* gene family. In G1 each gene *FhWRKY-3*, *FhWRKY-9*, *FhWRKY-17*, *FhWRKY-30*, *FhWRKY-32*, and *FhWRKY-39* contained two WRKY domains respectively (Figure 3A). In the group 2A, *FhWRKY-2* and *FhWRKY-14* showed a single WRKY domain and both have 6 to 8 motifs. The Group 2B, all the members of this group such as *FhWRKY-7*, *FhWRKY-11*, *FhWRKY-16*, *FhWRKY-18*, *FhWRKY-22*, and *FhWRKY-34*, contained single WRKY domain, but *FhWRKY-42* contained WRKY domain and additional Uso1\_p115\_C superfamily domain at C-terminal. Most of the members of Group 2B have ranged 4 to 9 motifs. In group 2C, all the members showed a single WRKY domain, have ranged 3 to 6 motifs but *FhWRKY-25* showed an additional AP2 domain at the C-terminal and a WRKY domain at N-terminal. Moreover, in group 2D, genes such as *FhWRKY-1*, *FhWRKY-13*, *FhWRKY-23*, and *FhWRKY-43* showed Plant\_znclust domain at C-terminal and WRKY domain at its N-terminal. However, only *FhWRKY-12*, and *FhWRKY-24* have lack of

Plant\_znclust domain, and all of the members contained 4 to 6 motifs ranged in group 2D. Further, the genes *FhWRKY-19*, *FhWRKY-28*, and *FhWRKY-45* found in group 2E contained a single WRKY domain and have ranged 3 to 4 motifs. In group G3, all the members also contained a single WRKY domain and had ranged of 4 to 6 motifs (Figure 3B).

#### Comparative phylogenetic analysis, chromosomal distribution, and gene duplications of *FhWRKY* genes

The highly conserved WRKYGQK motif logo was predicted in most of the *FhWRKY* genes (Figure 4A). However, the genes *FhWRKY4* and *FhWRKY15* have mutated heptapeptide sequence WRKYGKK at C-terminal (Supplemental File S2). To further explore the evolutionary relationship between *Fortunella hindsii* WRKY genes and *Arabidopsis*. We constructed phylogenetic tree by using the neighbor-joining method which spontaneously reflects the evolutionary status, homology, diversity, and grouping attribution of WRKY gene family (Figure 4B). The phylogenetic tree results showed that total 46 *FhWRKY* proteins based on the motifs and domains present in these peptide sequences can be divided into three main groups G1, G2, and G3. The members of G1 have two WRKY domain while other groups members have only a



**Figure 4.** Motif sequence logo, phylogenetic analysis, chromosomes distribution and gene duplications of *Fortunella hindsii* WRKYs: (A) the Motif sequence logos were based on the alignment of FhWRKY domains and (B) phylogenetic analysis between the *Fortunella hindsii* WRKY genes and *Arabidopsis*. The genes of *Fortunella hindsii* were named as FhWRKYs and *Arabidopsis* as AtWRKYs. The genes of FhWRKY were divided into seven subgroups. G1 is represented by blue color, 2A represented in grey color, 2B is shown in brown color, 2C represents in red color, 2D is represented in purple color, 2E is shown in yellow color and G3 is represented in green color. 95 WRKY genes were involved in this evolutionary analysis, which was carried out by using the NJ method with 1000 bootstraps. (C) Visualization of FhWRKY genes duplication across nine chromosomes. Vertical bars represent each chromosome number have gene names correspond to the approximate location of each WRKY-related gene.

single WRKY domain. The G1 group contained 15 proteins, of which six belonged to *Fortunella hindsii* and nine proteins from *Arabidopsis*. The group G2 was further divided into five sub-groups named 2A, 2B, 2C, 2D, and 2E based on the amino acid sequences similar to *Arabidopsis* proteins.<sup>38</sup> Further, 2A group had total five members, of which two belonged to FhWRKY and three AtWRKY proteins; 2B group contained total 12 members, of which seven FhWRKY and five AtWRKY proteins; the most prominent group was 2C, which contained total 29 members, out of which 16 were FhWRKY and 13 AtWRKY proteins; group 2D had total 12 members, which contained six FhWRKY and six AtWRKY proteins; 2E group had total six members, which contained three FhWRKY and three AtWRKY proteins. Furthermore, the G3 group had 16 members, which contained six FhWRKY and 10 AtWRKY proteins (Figure 4B).

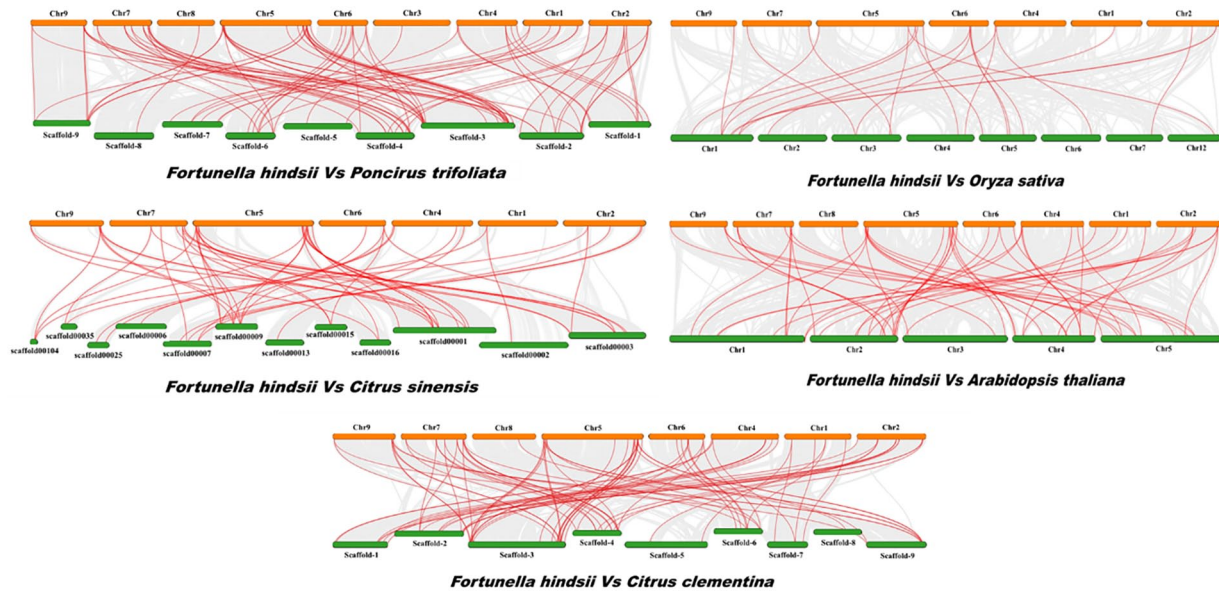
To investigate duplication events among FhWRKY genes in the *Fortunella hindsii* genome, a syntenic analysis was conducted. The results showed that certain FhWRKY genes, classified within the same subgroup based on phylogenetic analysis, are distributed across different chromosomes and linked through syntenic blocks. This indicates that segmental duplication might have contributed to the evolutionary expansion of the FhWRKY gene family (Figure 4C). All 46 FhWRKY genes

were mapped to nine chromosomes in the *Fortunella hindsii* genome. Particularly, although WRKY genes distribution across the chromosomes were not uniform, they were found on each chromosome. The maximum number of FhWRKY genes (12 and 8) were found on the Chr5 and Chr7, and the lowest number of FhWRKY genes (one) were found in Chr3 (Figure 4C). Different colors joining lines according to phylogenetic groups show the duplicated FhWRKY genes in the genome. Results indicated that 100% segmental duplication events may lead to FhWRKY genes and that these occurrences may have had a role in the evolution of FhWRKYs.

#### Orthologous gene pairs between *Fortunella hindsii* and dicot/monocot plants

The orthologous gene pairs provide information regarding evolutionary relationships between various plant species. In order to better understand the relationship of FhWRKY genes in dicot/monocot plant species, a dual synteny analysis was conducted between *Fortunella hindsii*, *Poncirus trifoliata*, *Citrus clementina*, *Citrus sinensis*, *Arabidopsis*, and *Oryza sativa*. Collinearity analysis results indicates that the number of orthologous events of FhWRKY-CcWRKY, FhWRKY-PtWRKY, FhWRKY-CsWRKY, and FhWRKY-AtWRKY show





**Figure 5.** Dual synteny analysis of *Fortunella hindsii* WRKY genes with dicot/monocot plants. Orange bars represent chromosomes of *Fortunella hindsii*, while green bars represent scaffolds/chromosomes of *Poncirus trifoliata*, *Citrus clementina*, *Citrus sinensis*, *Arabidopsis*, and *Oryza sativa*. Gray lines denote collinear blocks, while red lines emphasize syntenic WRKY gene pairs within the genomes of *Fortunella hindsii* and other plants.

higher divergence pattern, revealed that the divergence of the *Fortunella hindsii* and citrus plants happened subsequent to that of the common ancestor of dicotyledons. The high level of syntenic conservation indicated that WRKY TFs in citrus might have similar functions and structures to those of orthologs in *Poncirus trifoliata*, *Citrus clementina*, and *Citrus sinensis*. However, lower divergence pattern was observed in monocot *FhWRKY-OsWRKY* (Figure 5). The WRKY gene family's expansion is predominantly driven by gene duplication events, which are generally recognized as a crucial mechanism for increasing gene family members in plants. These duplications can be segmental, tandem, or whole-genome duplications, all of which are frequent in plant genomes. Gene duplication provides a foundation for evolutionary innovation, enabling functional diversification of duplicated genes. The WRKY family's expansion has led to the development of different roles in regulating plant responses to biotic and abiotic stressors, including pathogen defense, drought, and salinity tolerance.

#### Homology modeling of *FhWRKYs* proteins

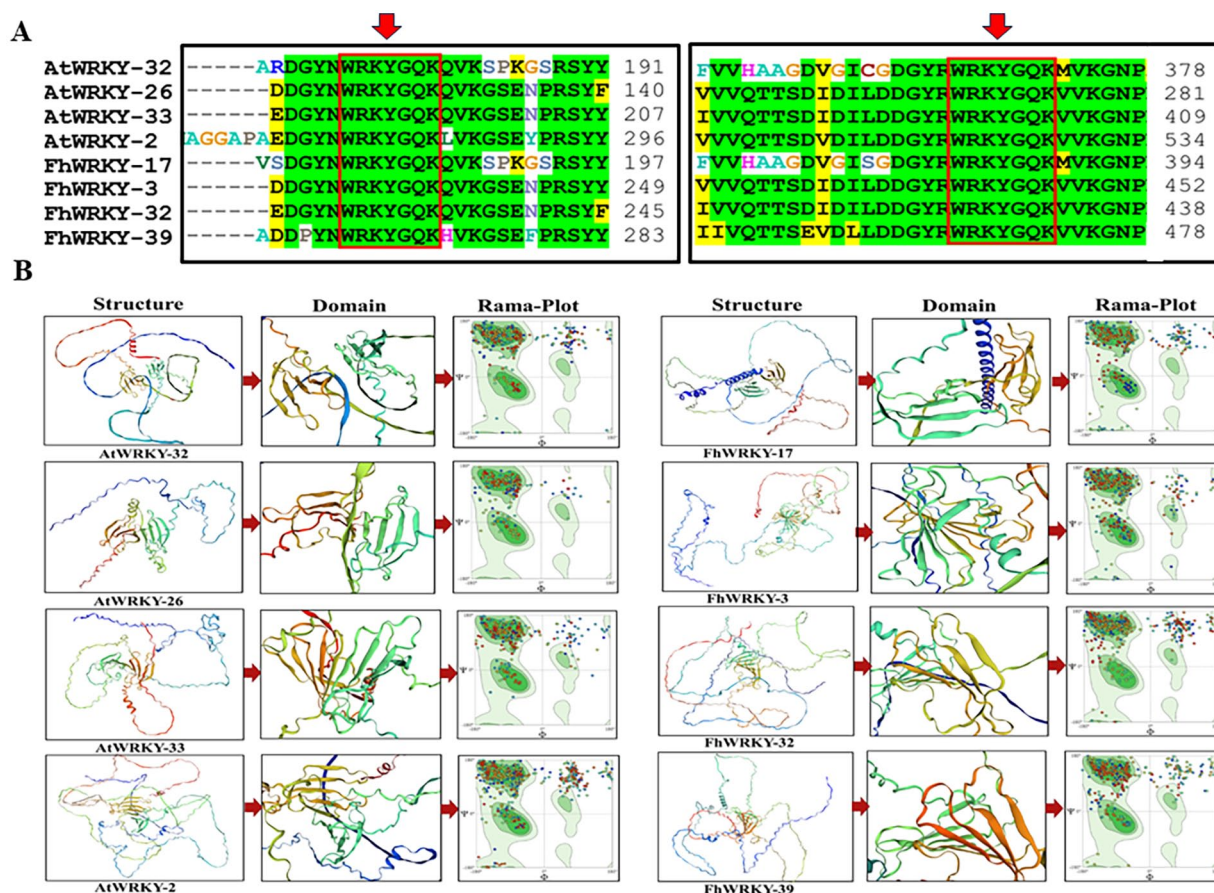
To determine the structure prediction of the proteins that revealed homology between them, we used SWISS-MODEL for 3D structure of FhWRKYs and AtWRKYs proteins. The members of group 1 were used to predict the homology modeling of FhWRKYs proteins. As mentioned above and according to the alignment of G1 proteins corresponding to *Arabidopsis* proteins contained two WRKY domains which contained conserved region “WRKYGQK” at specific positions in the protein sequences (Figure 6A).

To predict the three-dimension structure, a measuring parameter GMQE (Global Model Quality Estimate) was chosen for

model evaluation, which combines the features of target template alignment and structure with a score between 0 and 1. A higher score indicates greater reliability in the predicted structure quality. Additionally, the ERRAT program further assesses the model quality, with higher values reflecting better models in each evaluation criterion. In this analysis, we predicted the structure of proteins belongs to group one from the phylogenetic tree. The results are shown in (Figure 6B), which demonstrates the models of FhWRKYs corresponding to the AtWRKYs proteins in the same clade. The predicted structure of FhWRKYs closely resembled that of the corresponding AtWRKYs, indicating a higher degree of functional conservation. We found that all the structures of proteins showed complete homogeneity among members of the same group (Figure 6B). Further, MolProbity results and Ramachandran plot assess the conformational angles (psi) and (phi) of amino acid residues in protein structure and structure quality. Analysis revealed that 65% to 80% of the residues in our protein structures have dihedral angles (psi and phi) that fall within the favored region of the Ramachandran plot, while less than 20 % of the residues fall in the outlier region (Figure 6B).

#### Cis-regulatory analysis of the predicted FhWRKY genes

The *cis*-regulatory elements are DNA-binding motifs located at promoter region of genes that regulate the transcription.<sup>39</sup> In silico CREs analysis can be conducted to assess the potential function of several genes.<sup>40</sup> In *FhWRKY* genes many *cis*-regulatory elements were detected at the promoter region. Several phytohormone-responsive motifs including ABRE (Abscisic acid-responsive element), AuxRE, TGA-element, (Auxin-responsive element), GARE, P-box, TATC-box (Gibberellin-responsive element), and TCA-element (Salicylic acid-responsive



**Figure 6.** Sequence similarity and structure prediction of the FhWRKYs corresponding to AtWRKYs proteins belongs to same group: (A) the alignment of representative proteins highlighted WRKY conserved region in red box at corresponding positions and (B) structure prediction showed two WRKY domains represented in green and orange colors. Ramachandran plot illustrates dihedral angles (Psi and Phi) of amino acid residues in favored and outlier region.

element) were detected in the promoter region, suggesting that the expression of *FhWRKY* genes may be regulated by multiple phytohormones (Figure 7).

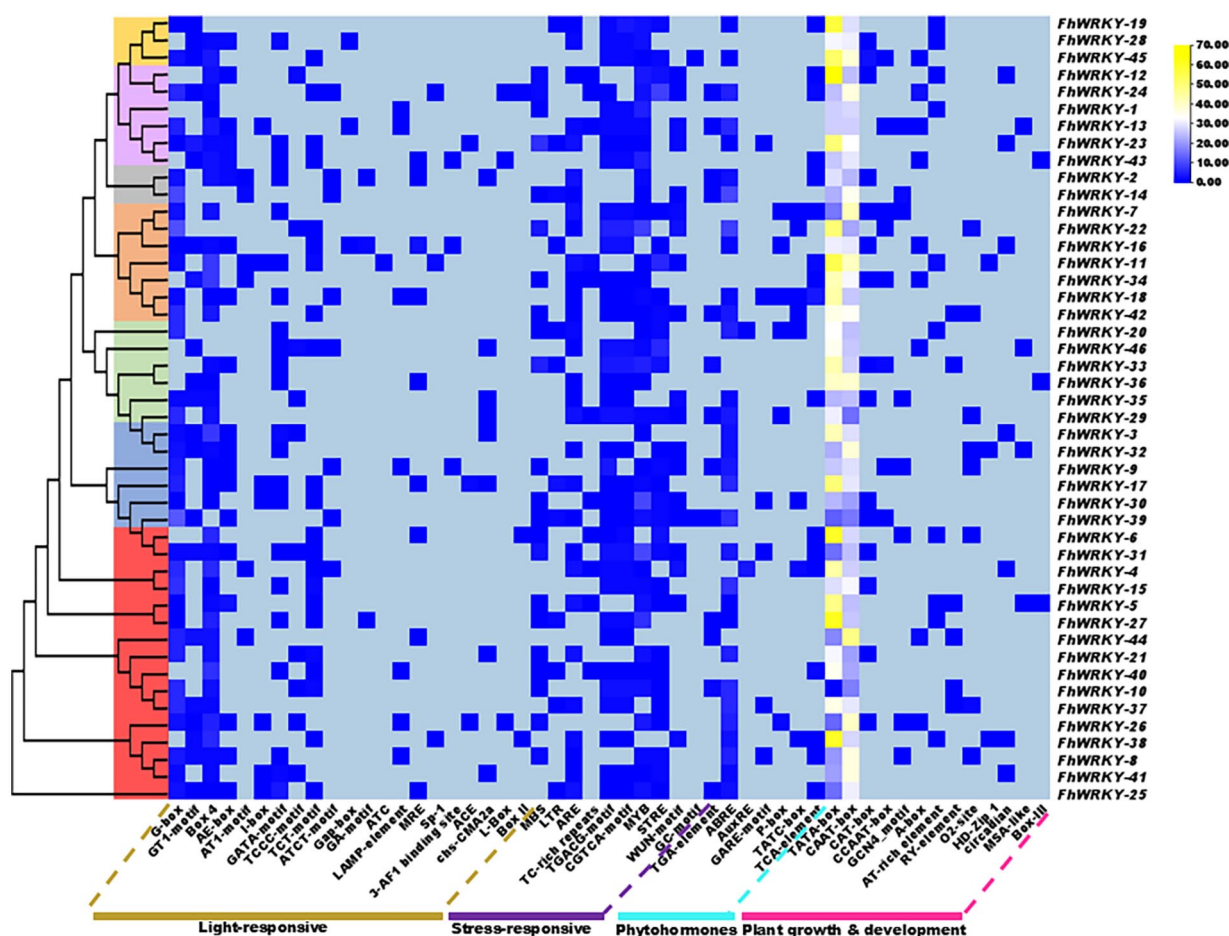
Moreover, stress related *cis*-regulatory elements motifs such as MBS, LTR, ARE, TC-rich repeats, TGACG-motif, CGTCA-motif, MYB, STRE, WUN-motif, and GC-motif were also found in the promoter region of *FhWRKY* genes (Figure 7). These findings indicated that *FhWRKY* genes may be closely associated with multiple biotic and abiotic stress responses. Furthermore, many motifs related to plant growth, development and other elements such as TATA-box, CAAT-box, CAT-box, CCAAT-box, GCN4-motif, A-box, AT-rich element, RY-element, O2-site, HD-Zip 1, circadian, MSA-like, and Box-III were detected in *FhWRKYs*. Several Light responsive elements such as G-Box, GT1-motif, Box 4, AE-box, AT1-motif, I-box, GATA-motif, TCCC-motif, TCT-motif, ATCT-motif, Gap-box, GA-motif, ATC, LAMP-element, MRE, Sp-1, 3-AF1 binding site, ACE, chs-CMA2a, L-Box, and Box II were also found in the promoter region of *FhWRKY* genes as shown in (Figure 7) and data are listed in Table S4 (provided in Supplemental File 1).

### Functional annotation of predicted FhWRKYs genes

To explore functional annotation, all the identified 46 *FhWRKY* genes were successfully annotated for their potential functions by using Egnog. The annotation results revealed that *FhWRKY* genes were characterized into two categories: molecular function and biological process (Figure 8). Gene ontology (GO) analysis showed that *FhWRKY* genes are significantly involved in several metabolic processes including primary metabolic processes regulation, cellular metabolic processes regulation, DNA-templated transcription regulation, DNA-templated transcription, regulation of nucleobase compound metabolic process, regulation of macromolecule metabolic process, regulation of gene expression, RNA metabolic process as illustrated in biological process category.

Furthermore, *FhWRKYs* are predicted to be more frequently involved in many molecular functions such as sequence-specific DNA binding, DNA-binding transcription factor activity, transcription regulator activity, DNA binding, nucleic acid binding, and organic cyclic compound binding (Figure 8). Additionally, the residual WRKYs were dispersed uniformly across the biological process and molecular function. In





**Figure 7.** In-silico analysis of *cis*-regulatory elements in the promoter region of the identified *FhWRKY* genes in their respective groups were associated with different phytohormone, responses to stress, growth and developmental processes, side bar represent the highest count based on the number of times a particular *cis*-element occurs within the promoter.

contrast, there were no genes involved in the cellular components category.

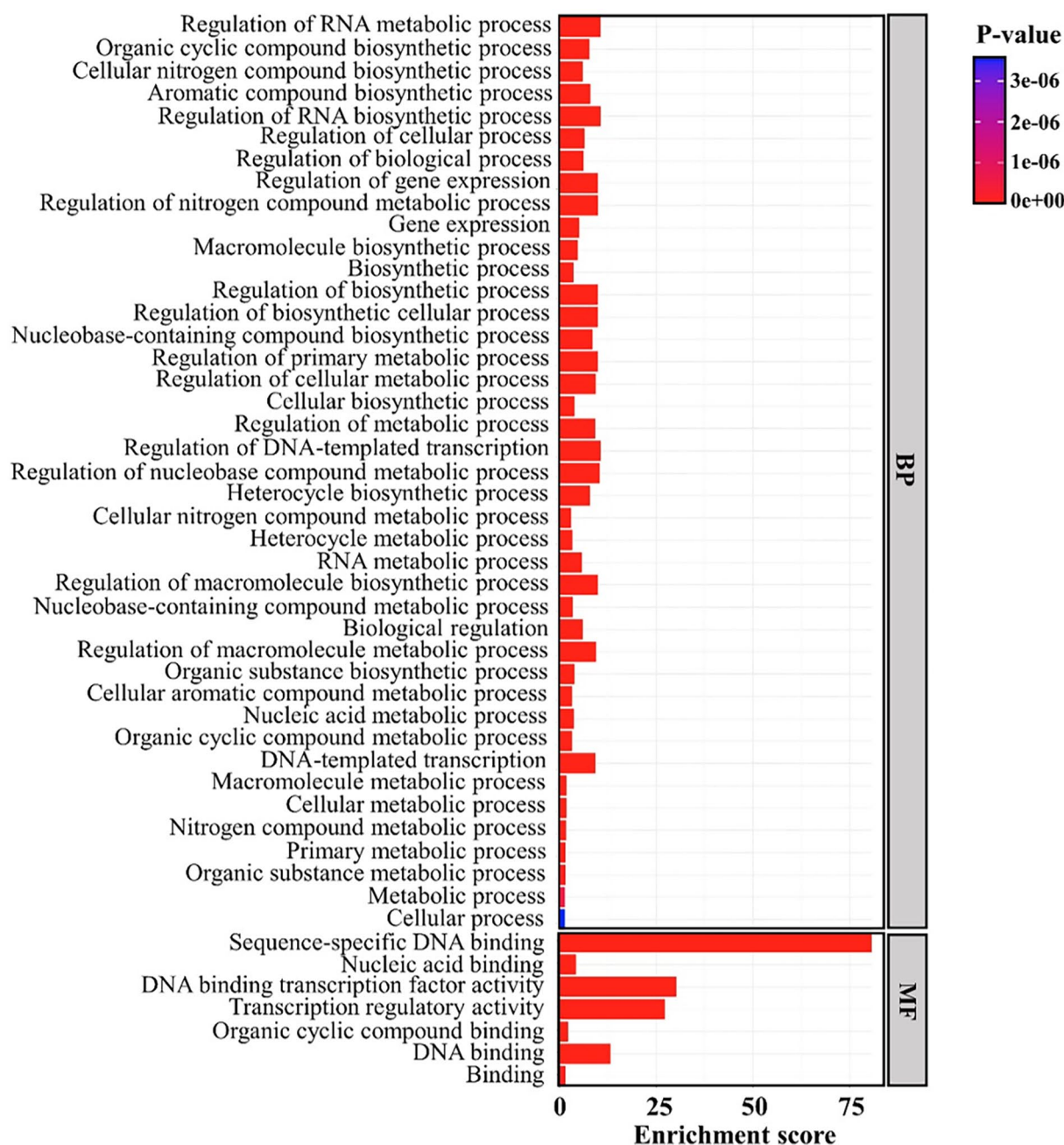
## Discussion

*Fortunella hindsii* is a mini citrus wild species characterized by short plant height and early flowering and is considered to be a potential model plant for citrus research.<sup>20</sup> The quality of fruit and production traits of *Fortunella hindsii* can be enhanced by conventional methods combined with molecular techniques and tools used through bioinformatic approaches. The WRKY protein family is a key transcription factor found in citrus and other organisms including monocots and dicots such as soybean,<sup>41</sup> rice,<sup>42</sup> wheat,<sup>43</sup> cotton,<sup>44</sup> and *Arabidopsis*.<sup>45</sup> So far, the whole genomes of numerous plants have been sequenced and a considerable number of *WRKY* genes have been found in various plant species.<sup>46,47</sup> A whole genome analysis of the *WRKY* genes is now feasible after the sequenced of the *Fortunella hindsii* genome.<sup>20</sup> However, there are no reports for the identification and analysis of the *WRKY* gene family have been published in *Fortunella hindsii*.

This one is the first study published publicly for the identification and analysis of *WRKY* transcription factors using

genomic data, resulting in a better understanding of the *WRKY* gene family in *Fortunella hindsii*. The number of transcription factors within a gene family is associated with the species genetic heritage and the long-term evolutionary series of plants. In this study, a total of 46 *FhWRKY* genes are identified in the *Fortunella hindsii* genome (Table S1). The number of identified *FhWRKY* genes was neither higher nor lower as compared to other plants, such as 79 *WRKY*s found in *Solanum tuberosum*, 57 *WRKY*s found in *Santalum album*, 71 *WRKY*s in *Panicum miliaceum* L., 74 *WRKY*s in *Arabidopsis thaliana*, 48 *WRKY*s in *citrus clementine*, 46 *WRKY*s in *citrus reticulata*, 1 *WRKY* in *citrus unshiu*, and 54 *WRKY*s in *Acer truncatum*,<sup>8,25,46,48-51</sup> suggesting that the *WRKY* gene family is highly conserved in *Fortunella hindsii*, which may be connected to gene duplication during the evolution and development of species.

The location of proteins is very important to determine their biological activity in the cells.<sup>52</sup> For evidence, most of the class I and II *NtWRKY* genes of common tobacco were localized in the nucleus.<sup>53</sup> Further evidence that most of the *Acer truncatum* *WRKY* genes for example *AtruWRKY1*, *AtruWRKY4*, and *AtruWRKY24* were found in the nucleus,



**Figure 8.** The gene ontology analysis of the identified *FhWRKY* genes. The lengths of the rectangular columns show the number of genes enrichment score that participated in the corresponding categories, significant to  $P$ -value  $< .05$ . BP means biological processes and MF mean molecular functions.

suggesting that the functions of *WRKY*s genes may be closely-related to the regulation the expression of target genes.<sup>25</sup> In this study, the results of In-silico sub-cellular localization showed that most of the putative *FhWRKY* proteins were located in the nucleus, except that few proteins *FhWRKY23*, *FhWRKY25*, *FhWRKY26*, and *FhWRKY27* were found in the peroxisome (Figure 1D).

Based on the sequence alignment and phylogenetics analysis *WRKY* genes are divided into three major groups such as GI, G-II, and G-III. This classification depends on the motif and conserved domain of *WRKY*. Previous studies reported that group G-I contains the highest number of *WRKY* genes in

*Populus trichocarpa*.<sup>54</sup> Whereas, group G-II was the largest group found in *Arabidopsis*, *Citrus reticulata*, and sesame (*Sesamum indicum* L.),<sup>45,46,55</sup> and G-III group contains the largest number of *WRKY* genes in rice.<sup>6</sup> However, we found that most of the *WRKY* genes in *Fortunella hindsii* are present in group G-II, indicating potential gene duplication during its evolutionary history. Therefore, our findings are consistent with the results of sesame and *Arabidopsis* but different from *Populus* and rice. Moreover, G-II are divided into five subgroups such as 2A, 2B, 2C, 2D, and 2E, according to the residues of amino acids. Subgroup 2C is also found the largest subgroup of G-II which is also consistent with the results of

*Arabidopsis* and rice<sup>42,43</sup>. There were six members of FhWRKY in each G-I and G-III, while the largest group was G-II containing 34 members by using WRKY proteins of *Arabidopsis* as reference respectively (Figure 4B).

Further, there is an evidence that double conserved WRKY domains have been found in G-I which can interact with the W-box core motif to activate downstream genes and G-II and G-III only had single WRKY domain and shared zinc finger motif (CX4-5CX22-23HXH or CX7Cx23HXC).<sup>4,45</sup> Another study reported that ScWRKY5 is a member of the group III WRKY family of the sugarcane (*Saccharum* spp.). It has only one zinc-finger motif (CX7Cx23HXC), which is the same reported as SIWRKY80 only one protein out of 81 WRKY proteins has this type of zinc-motif (CX7Cx23HXC) in tomato (*Solanum lycopersicum*), suggesting that this type of structural variant in zinc finger is unusual and may cause the presence or loss of certain functions.<sup>2,17</sup> In this study, we found that G-I group have two conserved WRKY domain; whereas, group G-II and G-III contained single WRKY domains (Figure 3). Our finding also indicated that groups G-I and G-II of *FhWRKY* genes shared zinc-finger motif (CX4-5CX22-23HXH), but group G-III of *FhWRKY* genes such as *FhWRKY33*, *FhWRKY35*, *FhWRKY36*, and *FhWRKY46* contained this type of zinc-finger motif (CX7Cx23HXC) (Table S3). The *WRKY* gene has a consistent sequence of heptapeptides WRKYGQK, it is unclear that adding homologous introns in distinct loci affects the functions of WRKYs proteins.<sup>4</sup> In this study we found highly conserved sequence region in most of the FhWRKYs, as shown in sequence logo (Figure 4A). Previous study reported that *citrus reticulata* *WRKY* gene such as *CrWRKY20* possess lack of heptapeptides sequence, mutations were observed at Q position, which is replaced by K, suggesting that during evolution *CrWRKY* genes may experience loss of WRKY domain.<sup>46</sup> In this study, our results showed that *FhWRKY4* and *FhWRKY15* have mutated heptapeptide sequence WRKYGKK at the C-terminal. Evolutionary changes have an important influence on the process of gene duplication. Tandem duplication refers to duplicated genes on the same chromosomes. However, a segmented mode of duplication occurs across different chromosomes.<sup>56</sup> The identified 46 WRKY transcription factors, which were unevenly distributed on nine chromosomes in *Fortunella hindsii*, showed segmental duplication and may play vital role in the evolution of WRKY family (Figure 4C). To further find the homologs of *Fortunella hindsii* in other plant species. Dual synteny blocks were constructed between *Fortunella hindsii* and monocot-dicot crops respectively, collinearity analysis showed that a higher number of orthologous events were observed among *Fortunella hindsii* with citrus plants and *Arabidopsis* (Figure 5).

The WRKY transcription factors may rapidly perceive signals of changes in external environment circumstances and transmit these signals promptly, allowing plants to respond to the environment.<sup>45</sup> In the signal transduction pathway of

phytohormones like ABA, stress-responsive genes can be activated by different abiotic stresses such as salt, heat, drought light etc. which increases the plant tolerance.<sup>57</sup> Moreover, to further know about the functions of WRKY proteins in *Fortunella hindsii*, *cis*-regulatory elements play an essential role in binding the TFs through their respective target site and promotes to regulate gene expression.<sup>25</sup> Previous studies reported that *cis*-regulatory elements of *CrWRKY* genes of *citrus reticulata* related to LTRE, LRE, and different phytohormones play a significant role in the plant growth and development.<sup>46</sup> In this study, we found that multiple copies of *cis*-regulatory related to phytohormones, stress, and light responsive elements were frequently found in the putative promoter region of *FhWRKY* genes, indicating that these genes may be closely related to multiple abiotic and biotic stress responses (Figure 7). By targeting these regulatory regions and manipulating the expression of stress-related *WRKY* genes, biotechnological technologies, such as CRISPR/Cas9 gene editing or genetic engineering could be applied to improve the tolerance of citrus crops to drought and diseases. This approach has already been successfully implemented in crops like rice, where *WRKY* genes were used to enhance stress resistance. As previously reported overexpression of particular *WRKY* genes in other species enhanced water-use efficiency and stress adaption, suggesting that exploiting such genes in citrus breeding programs could improve drought resistance.<sup>58</sup> Recent studies has improved our understanding of how WRKY transcription factors mediate plant responses to abiotic stresses such as drought, salt, and severe temperatures. For example, a study reported that numerous WRKY family members improved drought and heat tolerance in *Citrus sinensis* by regulating the ABA signaling pathway and ROS (reactive oxygen species) scavenging mechanisms.<sup>59</sup>

To explore functional annotation of predicted *FhWRKY* genes in *Fortunella hindsii*, showed that these genes were significantly involved in the DNA-binding transcription factor activity, nucleic acid binding, sequence-specific DNA binding, and RNA biosynthetic processes, and several metabolic processes (Figure 8). Although bioinformatics tools provide valuable insights into gene structure, phylogeny, and potential functions, biological experiments including gene expression profiling under various stress conditions or functional analysis through gene knockouts or overexpression in model organisms are necessary to confirm the roles of *WRKY* genes in stress tolerance or development.

## Conclusions

This is the first genome level study for the identification and analysis of *WRKY* gene family of *Fortunella hindsii*. We identified 46 *FhWRKY* genes, and all of them were unevenly distributed on nine chromosomes. Additionally, we also identified ten conserved domains and motifs of FhWRKY proteins, and



these *FhWRKY* genes based on the phylogenetic relationship with *Arabidopsis* were classified into three main groups G1, G2, and G3, while the G2 group was further divided into five sub-groups. Further, the collinearity analysis of *Fortunella hindsii* with citrus plants and *Arabidopsis* was better than with *Oryza sativa*. Sequence alignment and structure prediction analysis showed homology between WRKY *Fortunella* and *Arabidopsis* proteins. In silico CREs analysis showed that *FhWRKY* genes maybe closely related to multiple abiotic and biotic stress responses. Furthermore, the functional annotation result showed that most of the *FhWRKY* genes were involved in molecular functions and biological processes such as DNA-binding transcription factor activity, metabolic process, biological process cellular process, etc. The detailed computational analysis of FhWRKY proteins revealed in this study could be helpful for cloning and characterization at the molecular level, as well as gene expression and interaction studies with other transcription factors.

### Acknowledgements

The authors extend their appreciation for the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU250922].

### Author Contributions

HH and JMAK designed the research, methodology, data curation, writing – original draft preparation and project administration; HH, AA, IM, MN, OAD, BA analyze the data, writing – review & editing the manuscript; JMAK, funding acquisition. The authors read and approved the final manuscript.

### ORCID iD

Hadia Hussain  <https://orcid.org/0000-0001-7034-6574>

### Data Availability Statement

All the data generated or analyzed during this study are included in this published article and its supplementary data files. The *Fortunella hindsii* genome files v2.0 were retrieved from the Citrus Genome Database (CGD) (<https://www.citrusgenomedb.org/>). The Protein sequences of WRKY transcription factor gene families of *Arabidopsis thaliana* were obtained from TAIR (<https://www.arabidopsis.org/browse/genefamily/index.jsp>).

### Supplemental Material

Supplemental material for this article is available online.

### REFERENCES

- Haggag WM, Abouziena H, Kreem AEF, Habbasha ES. Agriculture biotechnology for management of multiple biotic and abiotic environmental stress in crops. *J Chem Pharm Res.* 2015;7:882-889.
- Huang S, Gao Y, Liu J, et al. Genome-wide analysis of WRKY transcription factors in *Solanum lycopersicum*. *Mol Genet Genom.* 2012;287:495-513.
- Ding ZJ, Yan JY, Li CX, Li GX, Wu YR, Zheng SJ. Transcription factor WRKY 46 modulates the development of *Arabidopsis* lateral roots in osmotic/salt stress conditions via regulation of ABA signaling and auxin homeostasis. *Plant J.* 2015;84:56-69.
- Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. *Trends Plant Sci.* 2010;15:247-258.
- Zhang Y, Wang L. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol Biol.* 2005;5:1-12.
- Wu KL, Guo ZJ, Wang HH, Li J. The WRKY family of transcription factors in rice and *Arabidopsis* and their origins. *DNA Res.* 2005;12:9-26.
- Wei KF, Chen J, Chen YF, Wu LJ, Xie DX. Molecular phylogenetic and expression analysis of the complete WRKY transcription factor family in maize. *DNA Res.* 2012;19:153-164.
- Ayadi M, Hanana M, Kharrat N, et al. The WRKY transcription factor family in citrus: valuable and useful candidate genes for citrus breeding. *Appl Biochem Biotechnol.* 2016;180:516-543.
- Ishiguro S, Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and  $\beta$ -amylase from sweet potato. *Mol Gen Genet.* 1994;244:563-571.
- Gu Y, Li W, Jiang H, et al. Differential expression of a WRKY gene between wild and cultivated soybeans correlates to seed size. *J Exp Bot.* 2017;68:2717-2729.
- Verweij W, Spelt CE, Blik M, et al. Functionally similar WRKY proteins regulate vacuolar acidification in petunia and hair development in *Arabidopsis*. *Plant Cell.* 2016;28:786-803.
- Lei Y, Sun Y, Wang B, et al. Woodland strawberry WRKY71 acts as a promoter of flowering via a transcriptional regulatory cascade. *Hort Res.* 2020;7:137.
- Lei R, Li X, Ma Z, Lv Y, Hu Y, Yu D. *Arabidopsis* WRKY 2 and WRKY 34 transcription factors interact with VQ20 protein to modulate pollen development and function. *Plant J.* 2017;91:962-976.
- Robatzek S, Somssich IE. Targets of AtWRKY6 regulation during plant senescence and pathogen defense. *Genes Dev.* 2002;16:1139-1149.
- Sun Y, Yu D. Activated expression of AtWRKY53 negatively regulates drought tolerance by mediating stomatal movement. *Plant Cell Rep.* 2015;34:1295-1306.
- Jin W, Zhou Q, Wei Y, et al. NtWRKY-R1, a novel transcription factor, integrates IAA and JA signal pathway under topping damage stress in *Nicotiana tabacum*. *Front Plant Sci.* 2018;8:2263.
- Wang D, Wang L, Su W, et al. A class III WRKY transcription factor in sugarcane was involved in biotic and abiotic stress responses. *Sci Rep.* 2020;10:20964.
- Yang Y, Chi Y, Wang Z, Zhou Y, Fan B, Chen Z. Functional analysis of structurally related soybean GmWRKY58 and GmWRKY76 in plant growth and development. *J Exp Bot.* 2016;67:4727-4742.
- Boonyaves K, Wu TY, Dong Y, Urano D. Interplay between ARABIDOPSIS GPB and WRKY transcription factors differentiates environmental stress responses. *Plant Physiol.* 2022;190:813-827.
- Zhu C, Zheng X, Huang Y, et al. Genome sequencing and CRISPR/Cas9 gene editing of an early flowering Mini-Citrus (*Fortunella hindsii*). *Plant Biotechnol J.* 2019;17:2199-2210.
- Yu C, Deng X, Chen C. Chromosomal characterization of a potential model mini-Citrus (*Fortunella hindsii*). *Tree Genet Genomes.* 2019;15:73.
- Yasuda K, Yahata M, Kunitake H. Phylogeny and classification of Kumquats (*Fortunella* spp.) inferred from CMA karyotype composition. *Hort J.* 2016;85:115-121.
- Dai WS, Peng T, Wang M, Liu JH. Genome-wide identification and comparative expression profiling of the WRKY transcription factor family in two Citrus species with different Candidatus *Liberibacter asiaticus* susceptibility. *BMC Plant Biol.* 2023;23:159.
- Wang N, Song X, Ye J, et al. Structural variation and parallel evolution of apomixis in citrus during domestication and diversification. *Natl Sci Rev.* 2022;9:114.
- Li Y, Li X, Wei J, et al. Genome-wide identification and analysis of the WRKY gene family and cold stress response in *Acer truncatum*. *Genes.* 2021;12:1867.
- Gasteiger E, Hoogland C, Gattiker A, et al. Protein analysis tools on the ExPASy server 571 571 from: the proteomics protocols handbook protein identification and analysis tools on the ExPASy server. In: Walker JM, ed. *The Proteomics Protocols Handbook*. Humana Press; 2019:571-607.
- Chen C, Wu Y, Li J, et al. TTools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol Plant.* 2023;16:1733-1742.
- Shui D, Sun J, Xiong Z, Zhang S, Shi J. Comparative identification of WRKY transcription factors and transcriptional response to *Ralstonia solanacearum* in tomato. *Gene.* 2024;912:148384.
- Liu C, Wang X, Xu Y, Deng X, Xu Q. Genome-wide analysis of the R2R3-MYB transcription factor gene family in sweet orange (*Citrus sinensis*). *Mol Biol Rep.* 2014;41:6769-6785.
- Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol.* 2021;38:3022-3027.
- Chow CN, Yang CW, Wu NY, et al. PlantPAN 4.0: Updated database for identifying conserved non-coding sequences and exploring dynamic transcriptional regulation in plant promoters. *Nucleic Acids Res.* 2024;52:1569-1578.

32. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. *Nucleic Acids Res.* 2015;43:39–49.
33. Wang Y, Li J, Paterson AH. MCScanX-transposed: detecting transposed gene duplications based on multiple colinearity scans. *Bioinformatics.* 2013; 29:1458–1460.
34. Chen C, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13:1194–1202.
35. Horton P, Park KJ, Obayashi T, et al. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 2007;35:585–587.
36. Sun S, Ma W, Mao P. Genomic identification and expression profiling of WRKY genes in alfalfa (*Medicago sativa*) elucidate their responsiveness to seed vigor. *BMC Plant Biol.* 2023;23:568.
37. Koralewski TE, Krutovsky KV. Evolution of exon-intron structure and alternative splicing. *PLoS one.* 2011;6: e18055.
38. Dong J, Chen C, Chen Z. Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. *Plant Mol Biol.* 2003;51:21–37.
39. Rozière J, Guichard C, Brunaud V, Martin ML, Coursol S. A comprehensive map of preferentially located motifs reveals distinct proximal cis-regulatory sequences in plants. *Front Plant Sci.* 2022;13:976371.
40. Jones DM, Vandepoel K. Identification and evolution of gene regulatory networks: insights from comparative studies in plants. *Curr Opin Plant Biol.* 2020;54:42–48.
41. Luo X, Bai X, Sun X, et al. Expression of wild soybean WRKY20 in Arabidopsis enhances drought tolerance and regulates ABA signalling. *J Exp Bot.* 2013;64:2155–2169.
42. Dai X, Wang Y, Zhang WH. OsWRKY74, a WRKY transcription factor, modulates tolerance to phosphate starvation in rice. *J Exp Bot.* 2016;67: 947–960.
43. Qin Y, Tian Y, Liu X. A wheat salinity-induced WRKY transcription factor TaWRKY93 confers multiple abiotic stress tolerance in Arabidopsis thaliana. *Biochem Biophys Res Commun.* 2015;464:428–433.
44. Dou L, Zhang X, Pang C, et al. Genome-wide analysis of the WRKY gene family in cotton. *Mol Genet Genom.* 2014;289:1103–1121.
45. Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 2000;5:199–206.
46. Maheen N, Shafiq M, Sadiq S, et al. Genome identification and characterization of WRKY transcription factor gene family in Mandarin (*Citrus reticulata*). *Agriculture.* 2023;13:1182.
47. Guo C, Guo R, Xu X, et al. Evolution and expression analysis of the grape (*Vitis vinifera* L.) WRKY gene family. *J Exp Bot.* 2014;65:1513–1528.
48. Zhang C, Wang D, Yang C, et al. Genome-wide identification of the potato WRKY transcription factor family. *PLoS One.* 2017;12:e0181573.
49. Yan H, Li M, Xiong Y, Wu J, Silva JAT, Ma G. Genome-wide characterization, expression profile analysis of WRKY family genes in *Santalum album* and functional identification of their role in abiotic stress. *Int J Mol Sci.* 2019;20:5676.
50. Yue H, Wang M, Liu S, Du X, Song W, Nie X. Transcriptome-wide identification and expression profiles of the WRKY transcription factor family in Broom-corn millet (*Panicum miliaceum* L.). *BMC Genom.* 2016;17:1–11.
51. Abdullah ZMR, Ahmad NNF, Govender N, Harun S, Mohd AN, Mohamed HZA. Comparative genome-wide analysis of WRKY, MADS-box and MYB transcription factor families in Arabidopsis and rice. *Sci Rep.* 2021;11:19678.
52. Silhavy TJ, Benson SA, Emr SD. Mechanisms of protein localization. *Microbiol Rev.* 1983;47:313–344.
53. Xiang X, Wu X, Chao J, et al. Genome-wide identification and expression analysis of the WRKY gene family in common tobacco (*Nicotiana tabacum* L.). *Hereditas.* 2016;38:840–856.
54. He H, Dong Q, Shao Y, et al. Genome-wide survey and characterization of the WRKY gene family in *Populus trichocarpa*. *Plant Cell Rep.* 2012;31:1199–1217.
55. Li D, Liu P, Yu J, et al. Genome-wide analysis of WRKY gene family in the sesame genome and identification of the WRKY genes involved in responses to abiotic stresses. *BMC Plant Biol.* 2017;17:1–19.
56. Moore RC, Purugganan MD. The evolutionary dynamics of plant duplicate genes. *Curr Opin Plant Biol.* 2005;8:122–128.
57. Hussain H, Cheng Y, Wang Y, et al. ASR1 and ASR2, two closely related ABA-induced serine-rich transcription repressors, function redundantly to regulate ABA responses in Arabidopsis. *Plants.* 2023;12:852.
58. Jiang Y, Liang G, Yu D. Activated expression of WRKY57 confers drought tolerance in Arabidopsis. *Mol Plant.* 2017;10:478–490.
59. Chen F, Hu Y, Vannozzi A, et al. The WRKY transcription factor family in citrus: mediators of abiotic stress response. *Plant Physiol.* 2022;188:208–221.