



Genome-Wide Identification and Comparative Analysis of WRKY Transcription Factors Related to Momilactone Biosynthesis in *Calohypnum plumiforme*

Yao Wang^{1,2†}, Ruiyan Zhu^{2†}, Min Shi², Qikai Huang², Siwei Zhang², Guoyin Kai^{2*} and Shuiliang Guo^{1*}

OPEN ACCESS

Edited by:

Liangsheng Zhang,
Zhejiang University, China

Reviewed by:

Chuan Yue,
Fujian Agriculture and Forestry
University, China
Youxiang Que,
Fujian Agriculture and Forestry
University, China
Jingya Zhao,
Shanghai Jiao Tong University, China

*Correspondence:

Guoyin Kai
guoyinkai@yahoo.com
Shuiliang Guo
guoshuiliang@163.com

†These authors share first authorship

Specialty section:

This article was submitted to
Phylogenetics, Phylogenomics,
and Systematics,
a section of the journal
Frontiers in Ecology and Evolution

Received: 05 November 2021

Accepted: 28 December 2021

Published: 02 February 2022

Citation:

Wang Y, Zhu R, Shi M, Huang Q,
Zhang S, Kai G and Guo S (2022)
Genome-Wide Identification
and Comparative Analysis of WRKY
Transcription Factors Related
to Momilactone Biosynthesis
in *Calohypnum plumiforme*.
Front. Ecol. Evol. 9:809729.
doi: 10.3389/fevo.2021.809729

¹ School of Life Sciences, Institute of Plant Biotechnology, Shanghai Normal University, Shanghai, China, ² Laboratory of Medicinal Plant Biotechnology, School of Pharmaceutical Sciences, Zhejiang Chinese Medical University, Hangzhou, China

Momilactones are diterpenoid phytoalexins with allelopathic functions, which have been found in the widely distributed bryophyte *Calohypnum plumiforme*. Clustered genes containing *CpDTC1/HpDTC1*, *CpCYP970A14*, *CpCYP964A1*, and *CpMAS* are involved in momilactone biosynthesis. Besides, momilactone concentration in *C. plumiforme* is affected by heavy metal treatment such as CuCl_2 . However, transcription factors which might regulate momilactone biosynthesis are unclear. WRKY transcription factors (TFs) regulate phytoalexin biosynthesis in many plant species. In this study, a systematic analysis of the WRKY TFs was performed according to the *C. plumiforme* genome. A total of 19 *CpWRKY* genes were identified and categorized into five subgroups based on their phylogenetic relationship. Conserved domain and motif analysis suggested that the WRKY domain was highly conserved, but there were some variations. *Cis*-acting elements and binding sites analysis implied that *CpWRKY* genes might be induced by stress and further regulate the biosynthesis of momilactones. Our study lays a foundation for further functional characterization of the candidate *CpWRKY* genes involved in the regulation of momilactone biosynthesis, and provides new strategies for increasing momilactone production.

Keywords: *Calohypnum plumiforme*, momilactones, heavy metal elicitation, WRKY, transcription regulation

INTRODUCTION

Calohypnum plumiforme (formerly *Hypnum plumaeforme*) is a widely distributed bryophyte in East Asia, which has been broadly used as a bio monitor to trace metals in the atmosphere (Okada et al., 2016; Li et al., 2020). Besides, *C. plumiforme* was used in traditional herb medicine to treat burns and epistaxis (Li et al., 2020). Recently, it was reported that the moss *C. plumiforme* owns allelopathic activity by suppressing the growth of plants around, mainly due to the production of chemical

Abbreviations: *C. plumiforme*, *Calohypnum plumiforme*; *Os*, *Oryza sativa*; TF, transcription factor; BGCs, biosynthetic gene clusters; GGPP, geranylgeranyl diphosphate; *DTC1*, syn-pimara-7, 15-diene synthase; *CYP*, Cytochrome P450 monooxygenase; *MAS*, momilactone A synthase.

defense compound momilactones (Kato-Noguchi and Kobayashi, 2009; Okada et al., 2016; Mao et al., 2020). Momilactone A and B were firstly found in the cultivated rice *Oryza sativa*, exhibiting remarkable antimicrobial and allelopathic functions (Kato-Noguchi et al., 2010; Toyomasu et al., 2014; Mao et al., 2020), which are biosynthesized through the biosynthetic gene clusters (BGCs) (Kato et al., 1973; Nützmann et al., 2018). Recently, the *C. plumiforme* genome was sequenced and annotated, with a genome size of 335 Mb (Mao et al., 2020). Notably, genome sequencing revealed a functional BGC responsible for momilactone biosynthesis in *C. plumiforme* (Mao et al., 2020).

In *C. plumiforme*, enzyme genes in the BGC containing *CpDTC1/HpDTC1*, *CpCYP970A14*, *CpCYP964A1*, and *CpMAS* have been well characterized recently (Okada et al., 2016; Mao et al., 2020). *HpDTC1* (encoding syn-pimara-7, 15-diene synthase) was verified as a bifunctional diterpenoid cyclase that catalyzed the two-step cyclization of the common precursor geranylgeranyl diphosphate (GGPP) to form syn-pimara-7, 15-diene (Okada et al., 2016). Then, two cytochrome P450, named *CYP970A14* and *CYP964A1*, oxidized C19 and C3 position of syn-pimara-7, 15-diene to form 3OH-syn-pimaradienolide (Mao et al., 2020). Subsequently, the momilactone A synthase *CpMAS* catalyzed 3OH-syn-pimaradienolide to form momilactones A (Mao et al., 2020). The momilactones BGC in rice is somewhat different from that in *C. plumiforme*. In rice, the production of syn-pimara-7, 15-diene synthase requires the catalysis of both *OsCPS4* and *OsKSL4* (Otomo et al., 2004; Wilderman et al., 2004). The transformation from syn-pimara-7, 15-diene to 3OH-syn-pimaradienolide requires the consecutive oxidation of at least three P450 (*CYP99A2/A3*, *CYP707A8*, and *CYP76M8*) (Wang et al., 2011; Kitaoka et al., 2015). *OsMS* has 52.3% identity to *CpMAS*, and responsible for converting 3OH-syn-pimaradienolide to momilactones A (Kitaoka et al., 2016). However, the gene encoding the enzyme that converts momilactones A to B in all momilactones-producing plants is still unclear.

Stress treatment is an efficient strategy to promote the production of momilactones A and B in *C. plumiforme*. For example, concentration of momilactone A and B in *C. plumiforme* in UV-irradiated condition was 14- and 15-fold greater than that of non-UV irradiated (Kato-Noguchi and Kobayashi, 2009). Jasmonic acid-treated *C. plumiforme* produced higher content of momilactone A and B than in non-treated group (Kato-Noguchi and Kobayashi, 2009). CuCl_2 -treated *C. plumiforme* produced more than 6-folds momilactone A and B than the control (Kato-Noguchi and Kobayashi, 2009), along with expression level of the four clustered genes much higher than the untreated samples (Mao et al., 2020). Previous studies have shown that different transcription factors (TFs) were induced under CuCl_2 treatment including WRKY TFs (Okada et al., 2016). Meanwhile, we found that variable W-box existed in the promoters of clustered genes.

WRKY is the largest transcription factor family in plants, which are involved in responses to biotic and abiotic stress (Eulgem et al., 2000; Rushton et al., 2010; Phukan et al., 2016). WRKY TFs is named because of its highly conserved

WRKY domain consist of N-terminal WRKYGQK domain and a C-terminal zinc finger structure (Eulgem et al., 2000; Eulgem and Somssich, 2007). According to the number of WRKY domains and the type of zinc finger structure, the WRKY family can be divided into three groups: Group I contains two WRKY domains and a C_2H_2 zinc finger structure; Group II contains a WRKY domain and a C_2H_2 type zinc finger structure; Group III harbors a WRKY domain and a C_2HC zinc finger structure (Zhang and Wang, 2005). Most WRKY TFs bind to W-box (TTGACT/C), which located in the promoters of many biosynthetic and defense-associated genes (Eulgem et al., 2000; Rushton et al., 2010). WRKY genes are involved in secondary metabolism and disease resistance (Ryu et al., 2006; Yokotani et al., 2013; Zhang R. et al., 2021). For example, *ZmWRKY79* promoted the accumulation of maize terpenoid phytoalexins through binding to biosynthetic gene promoters (Fu et al., 2018). *OsWRKY53* was phosphorylated by the *OsMKK4-OsMPK3/OsMPK6* cascade and positively regulated the diterpenoid phytoalexins (DP, momilactones and phytocassane, et al.) biosynthetic genes (Chujo et al., 2014). Overexpression of phosphomimic mutated *OsWRKY53* increased the resistance of rice to blast fungus (*Magnaporthe grisea*) (Chujo et al., 2014). Nevertheless, *OsWRKY76* is a negative transcription factor of DP biosynthesis, which playing opposite roles in blast disease resistance (Bagnaresi et al., 2012; Yokotani et al., 2013). Overexpression of *OsWRKY45* in rice plants showed extremely strong resistance to *Magnaporthe oryzae*, and further study revealed that the overexpression of *OsWRKY45* upregulated the DP biosynthetic genes such as *KSL4* and *CYP99A2* (Akagi et al., 2014).

Although WRKY transcription factors have been identified in many plant species, information on WRKY TFs in *C. plumiforme* has not been reported. In this study, WRKY proteins in bryophytes *C. plumiforme* were identified using bioinformatic methods based on the genomic information, and classification and evolutionary relationships, conserved motif and gene structure, *cis*-acting elements and binding sites were systematically analyzed. This study provides available data to reveal the biological function of *CpWRKY* proteins, and lays a theoretical foundation for the discovery of potential *CpWRKY* proteins that regulates the biosynthesis of momilactones.

MATERIALS AND METHODS

Gene Identification and Characteristics

The WRKY hidden Markov model was obtained from Pfam database (PF03106),¹ which was used to search all putative WRKY genes from the *C. plumiforme* genome database (GSA accession: PRJCA001833) using Hmsearch tools (Finn et al., 2011). Pfam (see text footnote 1), NCBI CDD,² and SMART³ were used to verify the presence of WRKY domain (Marchler-Bauer et al., 2015; Finn et al., 2016). Proteins without WRKY conserved

¹<http://pfam.xfam.org/>

²<http://www.ncbi.nlm.nih.gov/cdd>

³<https://www.ebi.ac.uk/EMBL/BLAST/>

domains were removed and 19 WRKY proteins were retained for further analysis. Additionally, the theoretical isoelectric point (pI) and molecular weight (MW) of the CpWRKY proteins were calculated using the ExPASy⁴ (Gasteiger et al., 2003). Wolf PSORT II⁵ was used to predict the subcellular localization of CpWRKY protein (Wang et al., 2021).

Multiple Sequence Alignment and Phylogenetic Analyses

WRKY protein sequences in *Arabidopsis thaliana* were downloaded from UniProt protein database,⁶ named as *AtWRKY1-AtWRKY 70* (Ding et al., 2020). Subsequently, the Clustal W multiple sequences alignment of CpWRKY (19) and *AtWRKY* (70) proteins was performed using MEGA6 software. The phylogenetic tree was constructed by using Neighbor-Joining method with bootstrap value as 1,000 (default for other parameters) (Tamura et al., 2013). The display of the phylogenetic tree was performed by Interactive Tree of Life (iTOL)⁷ (Letunic and Bork, 2007). A multiple sequence alignment of *AtWRKY* and CpWRKY proteins was performed with DNAMAN software.

Gene Structure and Motif Composition Analysis

Conserved motifs of CpWRKY proteins were analyzed through the MEME online tools,⁸ and the number of motifs was set to 10 (the rest of the reference default) (Bailey et al., 2006). Exon-intron structure of CpWRKY genes was determined by comparing the genome sequence with the corresponding coding sequence (CDS), and visualized by software TBtools (Chen et al., 2020).

Analysis of Promoter *cis*-Acting Elements

The upstream promoter region sequences of CpWRKY genes corresponding to 3,000 bp were extracted from the genomic database of *C. plumiforme*, and then submitted to PlantCARE⁹ for *cis*-acting elements annotation (Lescot et al., 2002). TBtools software was used to draw the location distribution map of related elements (Chen et al., 2020).

Analysis of Expression Pattern of CpWRKY Genes and Binding Sites in the Promoters of Momilactone-Biosynthesis Pathway Genes

Transcriptomic data (accession number PRJDB9779) of CuCl₂-treated *C. plumiforme* (which was treated with 500 μM CuCl₂ for 8 h) were downloaded from the public database of NCBI for further analysis. Heatmap of CpWRKYs and momilactone biosynthetic gene cluster including *CpDTC1*, *CpCYP964A1*, *CpCYP970A14* and *CpMAS* gene expression was constructed

⁴<http://web.expasy.org/protparam/>

⁵<https://psort.hgc.jp>

⁶<https://www.uniprot.org/>

⁷<http://itol.embl.de/>

⁸<http://meme.sdsc.edu/meme/itro.html>

⁹<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>

using TBtools software. *Cis*-elements involved in the promoter regions (3,000 bp) of these pathway genes were analyzed in PlantCARE (Lescot et al., 2002).

RESULTS

Identification of CpWRKY Genes in the *Calohypnum plumiforme* Genome

In this study, 19 CpWRKY genes were screened out in the complete *C. plumiforme* genome. The physical and chemical properties of CpWRKY genes were analyzed by ExPASy, and the results showed that the number of amino acids ranged from 385 aa (*CpWRKY15*) to 1043 aa (*CpWRKY9*). The molecular weight ranged from 42.6 kDa (*CpWRKY15*) to 111.8 kDa (*CpWRKY9*). The isoelectric point results showed that 12 CpWRKY members were less than 7, which were acidic protein, and 7 CpWRKY members were greater than 7, which were alkaline protein. The hydrophilic and hydrophobic results were all less than 0, indicating that CpWRKY members were hydrophilic proteins. The subcellular localization prediction results showed that 16 CpWRKY proteins were located in the nucleus (except *CpWRKY2*, *CpWRKY4* and *CpWRKY9*) (Table 1).

Phylogenetic Analysis and Classification of CpWRKY Genes

In order to investigate the phylogenetic and evolutionary relationship of WRKY genes in *C. plumiforme*, an unrooted phylogenetic tree was constructed (Figure 1). CpWRKY proteins were divided into five groups: IIa, IIb, IIc, IId, III. Notably, no WRKY proteins were found in group I and IIe. Group IIc has the most CpWRKY members, with high homology to *AtWRKY8*, *AtWRKY28*, *AtWRKY71*, *AtWRKY57*, *AtWRKY23*, *AtWRKY48* and *AtWRKY68*. Multiple sequence alignment revealed that all the CpWRKY proteins contained the conserved “WRKYGQK” domain and C₂H₂ or C₂HC Zinc finger structure. Additionally, the WRKY domain of subgroup III was in WKKYGNK form (Figure 2). The zinc finger structure at the C terminal was slightly different among the five groups, such as subgroup IIa contained CX₄CX₂₃HXH, while the zinc finger structure of subgroup III is CX₄CX₂₂HXC (Figure 2).

Conserved Motif Compositions and Gene Structure of CpWRKY Genes

MEME analysis revealed 10 conserved motifs in CpWRKY proteins (Figures 3A,B,D). Motif 1 and motif 2 are two relatively conserved motifs, which exist in all five subgroups. Nevertheless, CpWRKY proteins in some subgroups have their own unique conserved motifs. For example, motif 8 only appears in IIa CpWRKY, motif 9 is a conservative motif unique to IIc; motif 5 and motif 6 occur only in the IId (Figure 3B).

In addition, all 19 CpWRKY family members had exon-intron structure, and the number of exons was variable, which ranged from 3 to 13. Subgroup IIc and IId members have 3–4 exons (except *CpWRKY 9*, which has 7 exons), while the *CpWRKY* family members in groups IIa, IIb, and III have 8 or 13 exons at

TABLE 1 | Basic information of WRKY gene family in *Calohypnum plumiforme*.

Gene name	Type	WRKY domain	Zinc-finger	Size/aa	Molecular weight/Da	PI	GRAVY	Predicted subcellular localization
CpWRKY1	Ilc	WRKYGQK	CX4CX22HXXH	532	57494.76	6.07	-0.754	Nucleus
CpWRKY2	III	WRKYGNK	CX4CX22HXC	951	105335.67	6.81	-0.502	Chloroplast
CpWRKY3	Ilc	WRKYGQK	CX4CX22HXXH	529	58003.92	5.21	-0.732	Nucleus
CpWRKY4	Ilc	WRKYGQK	CX4CX22HXXH	578	62216.29	5.88	-0.836	Peroxisome
CpWRKY5	Ilc	WRKYGQK	CX4CX22HXXH	730	80816.86	6.27	-0.798	Nucleus
CpWRKY6	Ilb	WRKYGQK	CX5CX22HXXH	767	82758.43	6.68	-0.603	Nucleus
CpWRKY7	Ilc	WRKYGQK	CX4CX22HXXH	702	76747.23	5.96	-0.802	Nucleus
CpWRKY8	IId	WRKYGQK	CX5CX22HXXH	399	43846.20	9.77	-0.763	Nucleus
CpWRKY9	Ilc	WRKYGQK	CX4CX22HXXH	1,043	111837.98	8.29	-0.616	Chloroplast
CpWRKY10	Ilc	WRKYGQK	CX4CX22HXXH	495	55175.34	5.90	-0.798	Nucleus
CpWRKY11	Ilc	WRKYGQK	CX4CX22HXXH	511	55867.37	6.14	-0.762	Nucleus
CpWRKY12	Ilc	WRKYGQK	CX4CX22HXXH	516	56455.19	6.64	-0.790	Nucleus
CpWRKY13	III	WRKYGNK	CX5CX22HXXH	532	58692.87	4.90	-0.741	Nucleus
CpWRKY14	IId	WRKYGQK	CX5CX22HXXH	403	43983.29	9.65	-0.743	Nucleus
CpWRKY15	IId	WRKYGQK	CX5CX22HXXH	385	42618.03	9.58	-0.680	Nucleus
CpWRKY16	Ilb	WRKYGQK	CX5CX22HXXH	782	83366.76	8.01	-0.664	Nucleus
CpWRKY17	Ila	WRKYGQK	CX4CX23HXXH	627	67079.43	7.37	-0.918	Nucleus
CpWRKY18	Ila	WRKYGQK	CX4CX23HXXH	648	69348.10	7.68	-0.875	Nucleus
CpWRKY19	Ila	WRKYGQK	CX4CX23HXXH	936	99287.84	6.54	-0.540	Nucleus

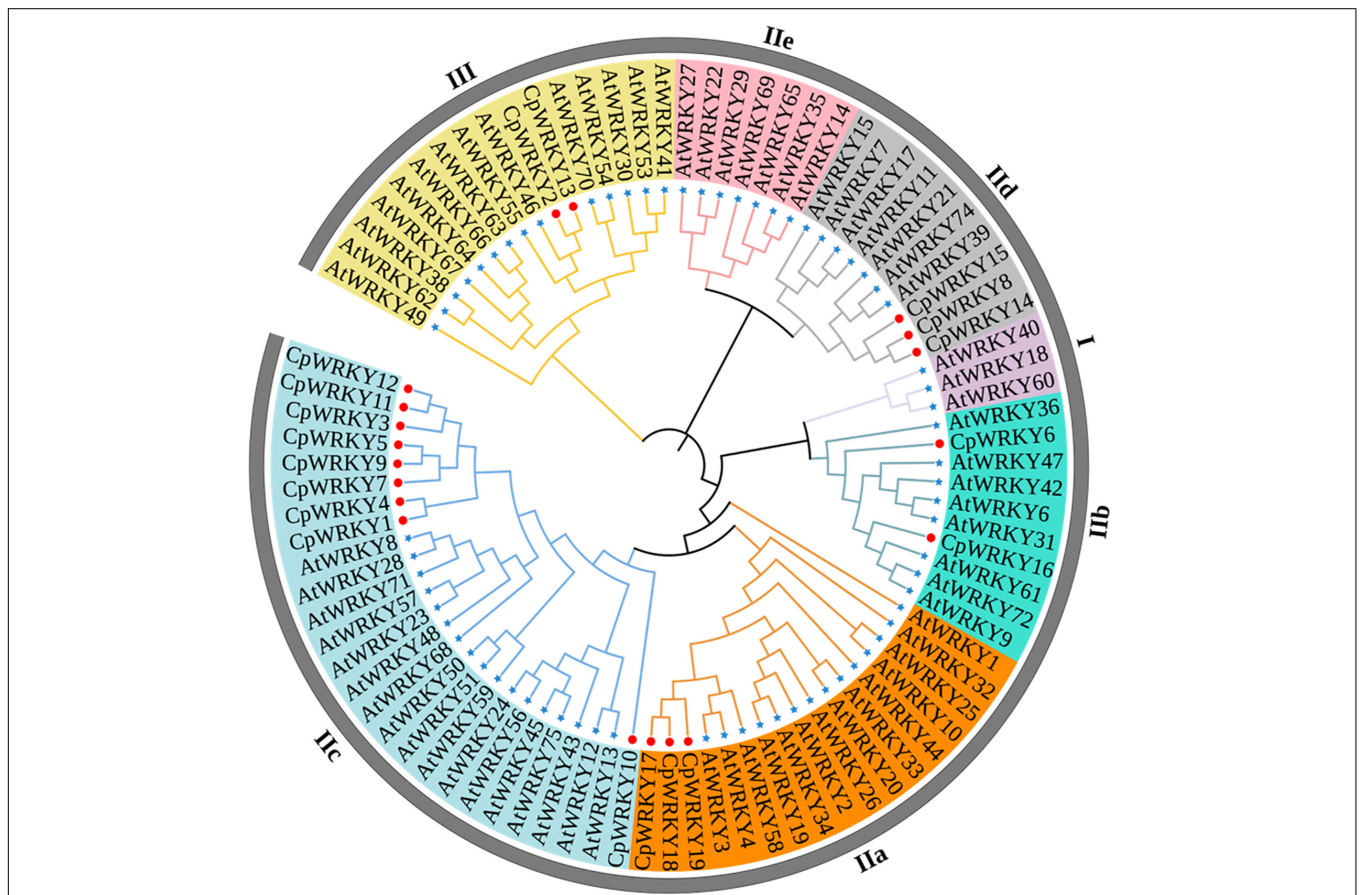


FIGURE 1 | Phylogenetic analysis of WRKY proteins in *Calohypnum plumiforme* and *Arabidopsis thaliana*. The WRKY domain protein sequences of *C. plumiforme* and *A. thaliana* were aligned with Clustal W, and the phylogenetic tree was constructed using MEGA 6 based on the neighbor-joining (NJ) method. The bootstrap value was 1,000 replicates. The proteins were classified into three groups and five subgroups. The colored branches indicate the different subgroups.

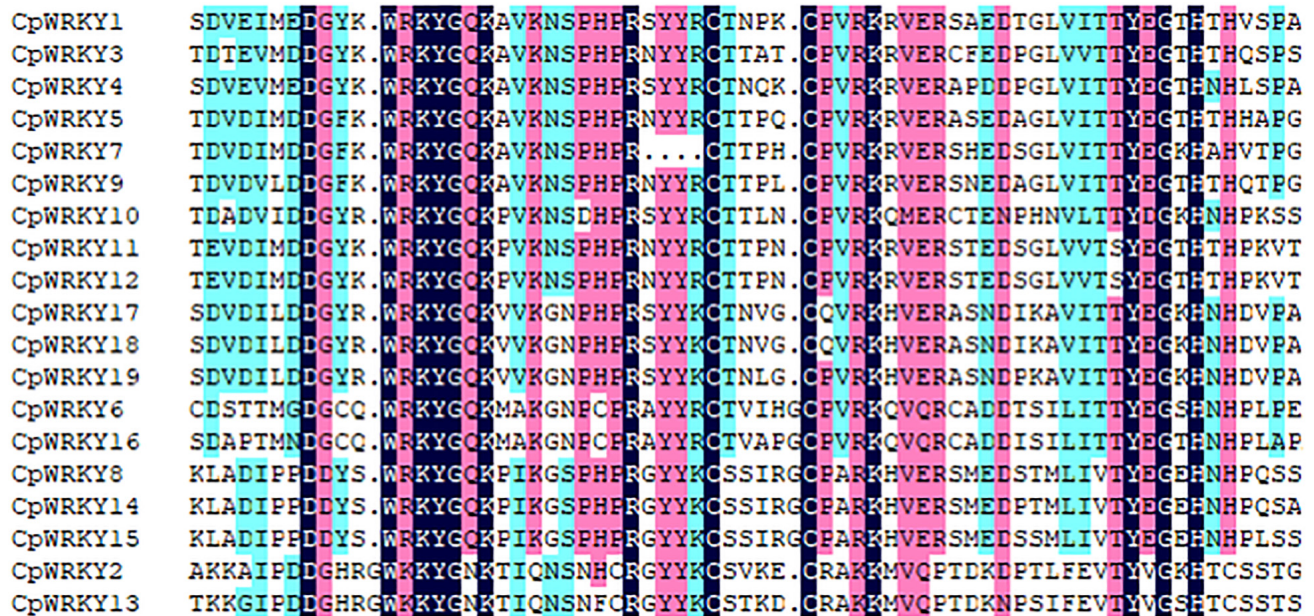


FIGURE 2 | Alignment of WRKY conserved domain sequences in *Calohyponum plumiforme*. The dark blue background represents conserved amino acid residues (100%), while the pink background represents the conserved amino acid residues (75%). Two red boxes represent the conserved heptapeptide domain and the variable zinc finger structure.

most (Figure 3C). These results showed that the structure of the CpWRKY gene family was significantly changed, indicating that the genome of *C. plumiforme* underwent large variation selection in the long-term evolution process.

Cis-Elements Analyses of CpWRKY Genes

In order to further investigate the function and regulatory mechanism of CpWRKY genes, we analyzed the *cis*-elements in promoters of CpWRKY genes (Figure 4). The results showed that CpWRKY gene promoters contained *cis*-elements related to plant growth and development, hormones and stress, among which hormones-related *cis*-elements were the most (236, accounting for 61.5% of the total). Among CpWRKY genes promoter sequences, MeJA-responsive element (CGTCA motif and TGACG motif) accounted for the largest proportion (120, 31.3%). It can be speculated that CpWRKYs play important and complex roles in plant response to hormone regulation (especially MeJA). CpWRKY7 had the most *cis*-elements (32) among all the CpWRKY genes, suggesting that it was relatively active and had more biological functions. Eighteen genes contained one or more ABRE (ABA-responsive element), TGACG-motif and CGTVA-motif (MeJA-responsive element), suggesting that these genes may be involved in the MeJA and ABA signal regulation pathway. Twelve genes contained one or more LTR (low temperature response element) and MBS (drought induced response element), indicating that these genes may be involved in the response to temperature stress. CpWRKY9 has a specific HD-ZIP1 (palisade mesophyll cells), which may play a specific role in leaf development (Figure 4).

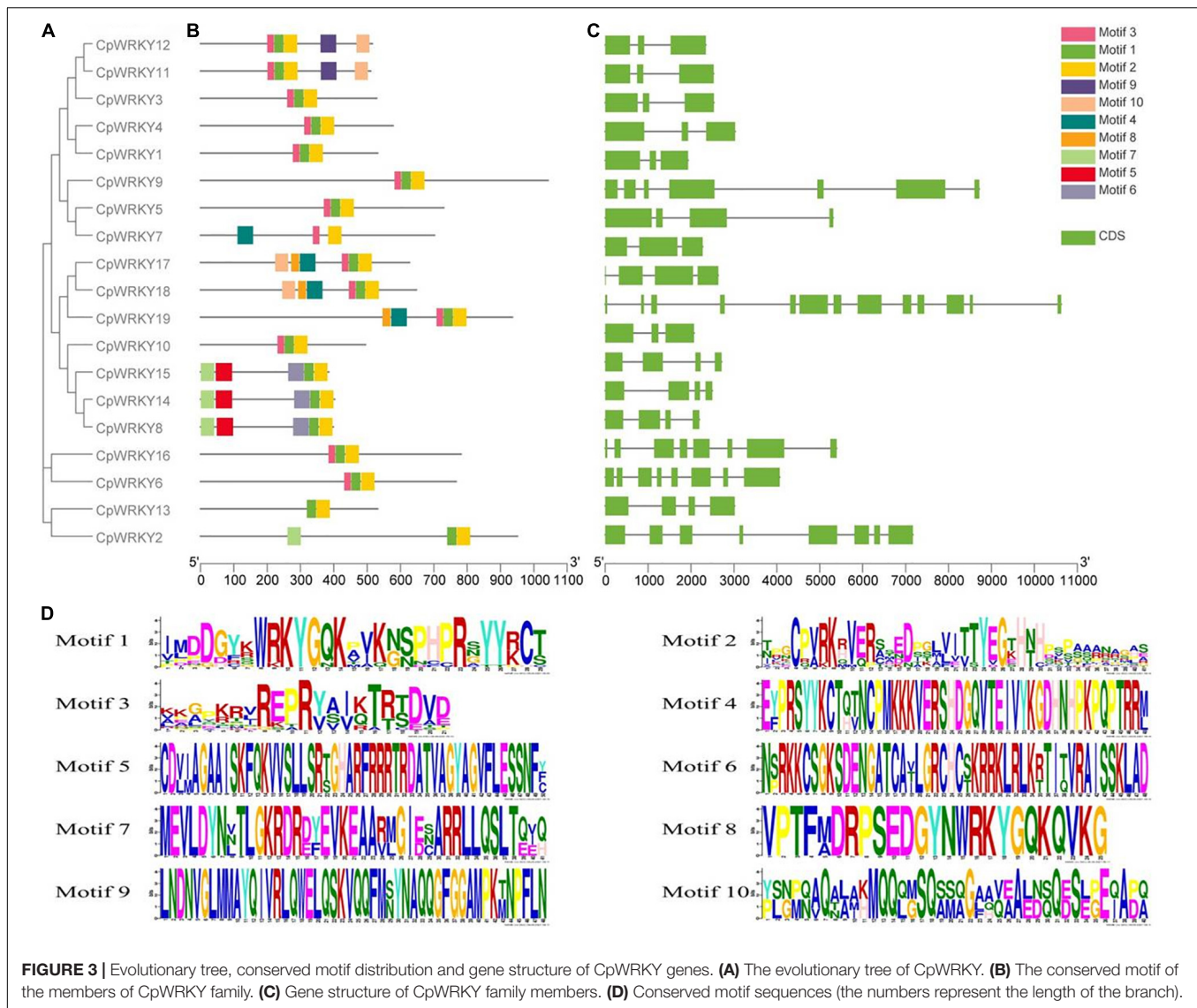
Expression Pattern of CpWRKY Genes and Binding Sites Analyses of Biosynthetic Gene Cluster

The expression patterns of BGCs and WRKY genes in the gametophytes of *C. plumiforme* treated with or without 500 μ M CuCl₂ for 8 h were analyzed. After 8 h treatment with CuCl₂, the expression levels of CpWRKY genes showed two patterns: CpWRKY1, CpWRKY3, CpWRKY4, CpWRKY5, CpWRKY6, CpWRKY7, CpWRKY9, CpWRKY10, CpWRKY11, CpWRKY12, CpWRKY13, CpWRKY17, CpWRKY18 and CpWRKY19 showed an upward trend, while CpWRKY2, CpWRKY8, CpWRKY14, CpWRKY15, and CpWRKY16 showed decreased expression (Supplementary Figure 1). Expression level of the clustered genes in momilactone biosynthesis was also elevated after CuCl₂ treatment (Supplementary Figure 2). Meanwhile, WRKY binding sites (W-box) were identified in the promoters of the BGCs. Two W-box elements existed in each of CpDTC1 and CpCYP970A14, four in CpCYP964A1 promoter, and five elements in CpMAS (Figure 5), suggesting that CpWRKY may play a regulatory role in momilactone biosynthesis.

DISCUSSION

Identification of CpWRKY Transcription Factors

WRKY TFs are a class of transcription factors that participate in many physiological and biochemical processes and exist widely in eukaryotes. With the development of high-throughput



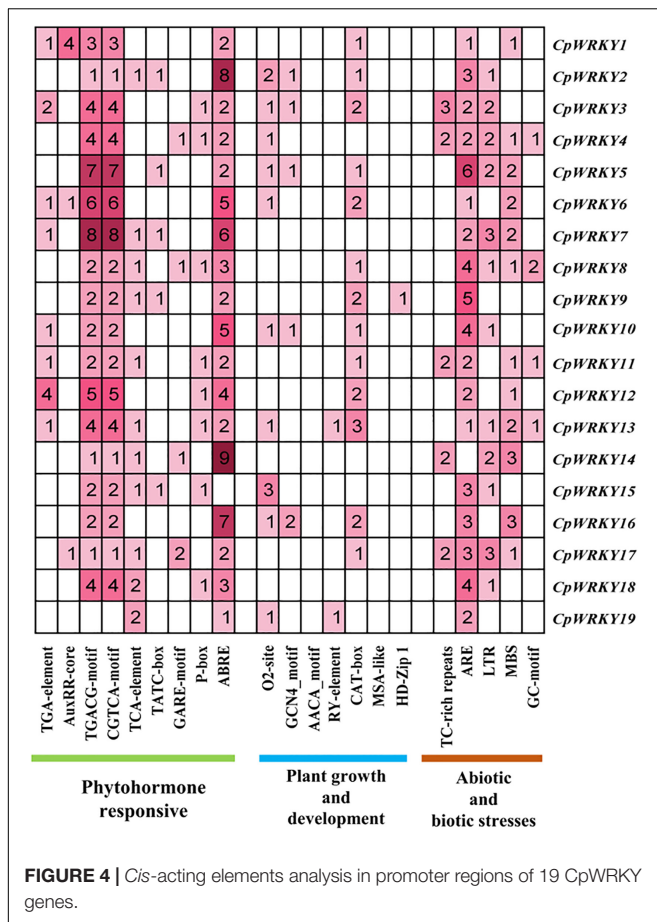
sequencing techniques, WRKY TFs have been identified in dozens of species (Di et al., 2021; Hao et al., 2021; Liu et al., 2021; Qu et al., 2021). Recent studies have shown that WRKY transcription factors are involved in the regulation of natural product biosynthesis (Yamada et al., 2021; Zhang L. et al., 2021). In this study, 19 WRKY TFs were excavated by analyzing the *C. plumiforme* genome. The number of WRKY transcription factors in *C. plumiforme* is much less than that of other higher plants, and the huge difference in the number is not only related to the genome of the species, but also related to the environmental influence of plants in the long-term evolution process (Rinerson et al., 2015).

Characterization of CpWRKY Transcription Factors

Phylogenetic analysis showed that *CpWRKY* proteins were divided into five groups: IIa, IIb, IIc, IId, III. Among the

five groups, the IIc harbored the most *CpWRKY* proteins, which was consistent with the most WRKY members of IIc subgroup in *A. thaliana*. However, there was no WRKY protein in group I in *C. plumiforme* genome, which was different from higher plants. In some higher plants, such as *Morus alba*, *Hordeum vulgare*, and *Sorghum bicolor*, the C-terminal domain of WRKY proteins classified into the group I may be the source of all WRKY proteins (Rinerson et al., 2015; Liu et al., 2017; Mao et al., 2019; Zhao et al., 2020). The WRKY proteins in the *C. plumiforme* do not seem to originate from this source.

The results of multiple sequence alignment showed that the WRKY domain of the two proteins in group III was mutated (“R” mutated to “K,” and “Q” mutated to “N”), which may affect the normal interaction between WRKY gene and W-box in downstream target genes, but it is also possible to gain the ability to identify other new motif, thereby obtaining new function (Zhou et al., 2008). Besides, MEME analysis showed that motif



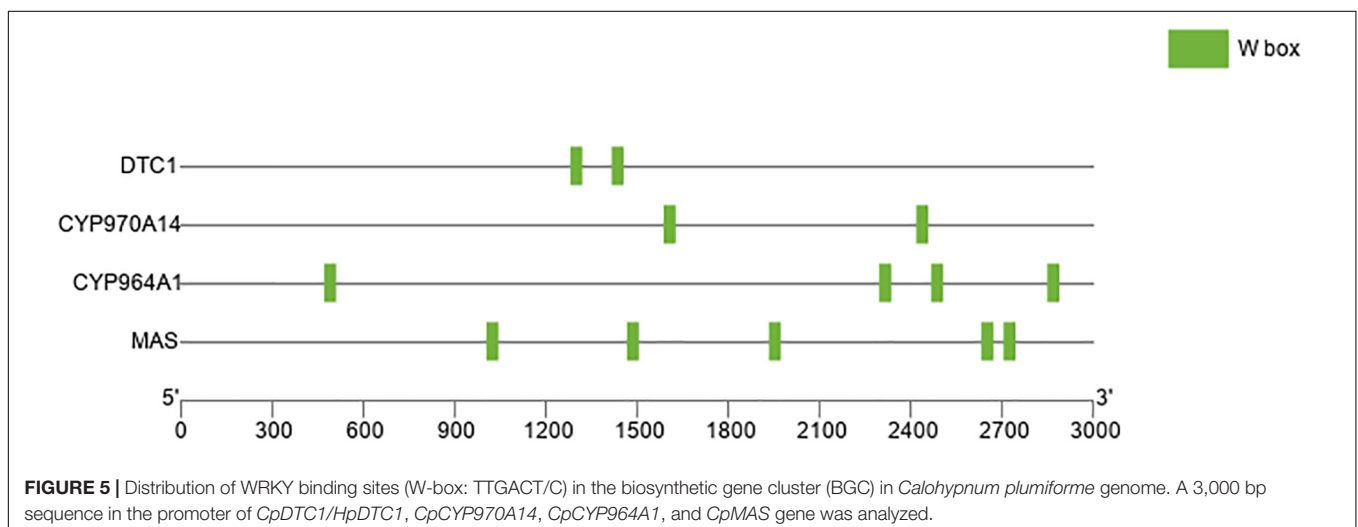
1 and motif 2 was highly conserved, which may be part of the WRKY conserved domain. Motif 8 and motif 9 were two relatively specific motifs, which may enable some WRKY proteins to have specific functions. Gene structure analysis showed that 3–13 exons were found in the genetic structure of *CpWRKY* genes, which provides available information for the study of the

evolutionary relationship of WRKYs between *C. plumiforme* and other plant species.

Possible Regulatory Function and Mechanism of *CpWRKY* Transcription Factors in Momilactone Biosynthesis

Previous studies have shown that the expression of momilactones biosynthesis genes in *O. sativa* and *C. plumiforme* can be induced by biotic and abiotic stress (Tamogami and Kodama, 2000; Rodrigues et al., 2004; Bi et al., 2007; Kato-Noguchi et al., 2007; Daw et al., 2008; Kato-Noguchi and Kobayashi, 2009; Kanno et al., 2012; Okada et al., 2016). CuCl₂ treatment increased the contents of momilactones and the expression level of momilactone biosynthesis genes. It is noteworthy that expression level of some CpWRKY genes was increased under CuCl₂ treatment (Okada et al., 2016). The analysis of cis-acting elements can provide data basis to analyze the relationship between stress and WRKY proteins (Li and Chen, 2015). CpWRKY family members contained cis-acting elements related to hormones and stresses, among which MeJA-responsive element accounted for the largest proportion, MeJA elicitation was efficient to enhance the momilactone A and B, and the exact function of CpWRKY on these clustered genes needs further study.

WRKY proteins specifically bind to the smallest motif TTGAC (C/T) in DNA called W-box (Eulgem et al., 2000). Most of the target gene promoters of WRKY TFs contain variable numbers of W-box, which are arranged in the same direction or form palindromic structures. WRKY TFs bind to them and regulate the expression of downstream functional biosynthetic genes or other transcription factors (Eulgem et al., 2000). Different amount of W-box was present in the BGCs genes (*CpDTC1*, *CpCYP964A1*, *CpCYP970A14*, and *CpMAS*). Given all that information, it is likely that CpWRKY TFs actively respond to CuCl₂ stress and stimulates the expression of momilactones biosynthetic genes by binding to W-box in the promoter regions



of target biosynthetic genes, thereby increasing the content of momilactones.

CONCLUSION

In this study, 19 WRKY family members were identified from the genome of the *C. plumiforme*, and their phylogenetic evolution, conserved domains, gene structure and conserved motif, *cis*-acting elements and binding sites were analyzed. These results provide valuable information for further elucidating the role of WRKY transcription factors in momilactones biosynthesis in *C. plumiforme*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

GK and SG conceived and designed the project. YW, RZ, QH, and SZ analyzed the data. YW, RZ, and MS wrote the manuscript.

REFERENCES

- Akagi, A., Fukushima, S., Okada, K., Jiang, C. J., Yoshida, R., Nakayama, A., et al. (2014). WRKY45-dependent priming of diterpenoid phytoalexin biosynthesis in rice and the role of cytokinin in triggering the reaction. *Plant Mol. Biol.* 86, 171–183. doi: 10.1007/s11103-014-0221-x
- Bagnaresi, P., Biselli, C., Orrù, L., Urso, S., Crispino, L., Abbruscato, P., et al. (2012). Comparative transcriptome profiling of the early response to *Magnaporthe oryzae* in durable resistant vs susceptible rice (*Oryza sativa* L.) genotypes. *PLoS One* 7:e51609. doi: 10.1371/journal.pone.0051609
- Bailey, T. L., Williams, N., Mislé, C., and Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34, W369–W373. doi: 10.1093/nar/gkl198
- Bi, H. H., Zeng, R. S., Su, L. M., An, M., and Luo, S. M. (2007). Rice allelopathy induced by methyl jasmonate and methyl salicylate. *J. Chem. Ecol.* 33, 1089–1103. doi: 10.1007/s10886-007-9286-1
- Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009
- Chujo, T., Miyamoto, K., Ogawa, S., Masuda, Y., Shimizu, T., Kishi-Kaboshi, M., et al. (2014). Overexpression of phosphomimic mutated OsWRKY53 leads to enhanced blast resistance in rice. *PLoS One* 9:e98737. doi: 10.1371/journal.pone.0098737
- Daw, B. D., Zhang, L. H., and Wang, Z. Z. (2008). Salicylic acid enhances antifungal resistance to *Magnaporthe grisea* in rice plants. *Aust. Plant Pathol.* 37, 637–644. doi: 10.1071/AP08054
- Di, P., Wang, P., Yan, M., Han, P., Huang, X., Yin, L., et al. (2021). Genome-wide characterization and analysis of WRKY transcription factors in *Panax ginseng*. *BMC Genomics* 22:834. doi: 10.1186/s12864-021-08145-5
- Ding, W., Ouyang, Q., Li, Y., Shi, T., Li, L., Yang, X., et al. (2020). Genome-wide investigation of WRKY transcription factors in sweet osmanthus and their potential regulation of aroma synthesis. *Tree Physiol.* 40, 557–572. doi: 10.1093/treephys/tpz129

YW, MS, GK, and SG revised the manuscript. All authors approved the final manuscript.

FUNDING

This work was supported by the Zhejiang Provincial Ten Thousand Program for Leading Talents of Science and Technology Innovation (2018R52050), the National Natural Science Fund of China (82073963, 31571735, 81522049), the Major Science and Technology Projects of Breeding New Varieties of Agriculture in Zhejiang Province (2021C02074), the Zhejiang Provincial Program for the Cultivation of High-level Innovative Health Talents, the Zhejiang Chinese Medical University Research Foundation (2021JKZDZC06), the Opening Project of Zhejiang Provincial Preponderant and Characteristic Subject of Key University (Traditional Chinese Pharmacology), and the Zhejiang Chinese Medical University (ZYAOX2018009).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2021.809729/full#supplementary-material>

- Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5, 199–206. doi: 10.1016/s1360-1385(00)01600-9
- Eulgem, T., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr. Opin. Plant Biol.* 10, 366–371. doi: 10.1016/j.pbi.2007.04.020
- Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367
- Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344
- Fu, J., Liu, Q., Wang, C., Liang, J., Liu, L., and Wang, Q. (2018). ZmWRKY79 positively regulates maize phytoalexin biosynthetic gene expression and is involved in stress response. *J. Exp. Bot.* 69, 497–510. doi: 10.1093/jxb/erx436
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563
- Hao, F., Yang, G., Zhou, H., Yao, J., Liu, D., Zhao, P., et al. (2021). Genome-wide identification and transcriptional expression profiles of transcription factor WRKY in Common Walnut (*Juglans regia* L.). *Genes (Basel)* 12:1444. doi: 10.3390/genes12091444
- Kanno, H., Hasegawa, M., and Kodama, O. (2012). Accumulation of salicylic acid, jasmonic acid and phytoalexins in rice, *Oryza sativa*, infested by the white-backed planthopper, *Sogatella furcifera* (Hemiptera: Delphacidae). *Appl. Entomol. Zool.* 47, 27–34. doi: 10.1007/s13355-011-0085-3
- Kato, T., Kabuto, C., Sasaki, N., Tsunagawa, M., Aizawa, H., Fujita, K., et al. (1973). Momilactones, growth inhibitors from rice, *Oryza sativa* L. *Tetrahedron Lett.* 14, 3861–3864. doi: 10.1016/s0040-4039(01)87058-1
- Kato-Noguchi, H., Hasegawa, M., Ino, T., Ota, K., and Kujime, H. (2010). Contribution of momilactone A and B to rice allelopathy. *J. Plant Physiol.* 167, 787–791. doi: 10.1016/j.jplph.2010.01.014
- Kato-Noguchi, H., and Kobayashi, K. (2009). Jasmonic acid, protein phosphatase inhibitor, metals and UV-irradiation increased momilactone A and B

- concentrations in the moss *Hypnum plumaeforme*. *J. Plant Physiol.* 166, 1118–1122. doi: 10.1016/j.jplph.2008.12.012
- Kato-Noguchi, H., Kujime, H., and Ino, T. (2007). UV-induced momilactone B accumulation in rice rhizosphere. *J. Plant Physiol.* 164, 1548–1551. doi: 10.1016/j.jplph.2006.12.008
- Kitaoka, N., Wu, Y., Xu, M., and Peters, R. J. (2015). Optimization of recombinant expression enables discovery of novel cytochrome P450 activity in rice diterpenoid biosynthesis. *Appl. Microbiol. Biotechnol.* 99, 7549–7558. doi: 10.1007/s00253-015-6496-2
- Kitaoka, N., Wu, Y., Zi, J., and Peters, R. J. (2016). Investigating inducible short-chain alcohol dehydrogenases/reductases clarifies rice oryzalexin biosynthesis. *Plant J.* 88, 271–279. doi: 10.1111/tpj.13249
- Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325
- Letunic, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127–128. doi: 10.1093/bioinformatics/btl529
- Li, J. L., Wie, L. L., Chen, C., Liu, D., Gu, Y. Q., Duan-Mu, J. X., et al. (2020). Bioactive constituents from the bryophyta *Hypnum plumaeforme*. *Chem. Biodivers.* 17:e2000552. doi: 10.1002/cbdv.202000552
- Li, Z. X., and Chen, X. B. (2015). Research advances on plant inducible promoters and related cis-acting elements. *Biotechnol. Bull.* 10, 8–15.
- Liu, C., Han, L. H., Song, P. B., Wang, D. Q., Wang, H. B., and Tang, L. Z. (2017). Genome-wide identification and bioinformatics analysis for mulberry WRKY transcription factors. *J. South. Agric.* 09, 1691–1699.
- Liu, J., Wang, X., Chen, Y., Liu, Y., Wu, Y., Ren, S., et al. (2021). Identification, evolution and expression analysis of WRKY gene family in *Eucommia ulmoides*. *Genomics* 113, 3294–3309. doi: 10.1016/j.ygeno.2021.05.011
- Mao, C. Z., Wang, L., Kong, D. D., Wang, H. D., Zhang, H. G., Sheng, Y. H., et al. (2019). Bioinformatics analysis of WRKY gene family in barley genome. *Mol. Plant Breed.* 23, 7623–7633. doi: 10.13271/j.mpb.017.007623
- Mao, L., Kawaide, H., Higuchi, T., Chen, M., Miyamoto, K., Hirata, Y., et al. (2020). Genomic evidence for convergent evolution of gene clusters for momilactone biosynthesis in land plants. *Proc. Natl. Acad. Sci. U.S.A.* 117, 12472–12480. doi: 10.1073/pnas.1914373117
- Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., et al. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi: 10.1093/nar/gku1221
- Nützmann, H. W., Scazzocchio, C., and Osbourn, A. (2018). Metabolic gene clusters in eukaryotes. *Annu. Rev. Genet.* 52, 159–183. doi: 10.1146/annurev-genet-120417-031237
- Okada, K., Kawaide, H., Miyamoto, K., Miyazaki, S., Kainuma, R., Kimura, H., et al. (2016). HpDTC1, a stress-inducible bifunctional diterpene cyclase involved in momilactone biosynthesis, functions in chemical defence in the moss *Hypnum plumaeforme*. *Sci. Rep.* 6:25316. doi: 10.1038/srep25316
- Otomo, K., Kenmoku, H., Oikawa, H., König, W. A., Toshima, H., Mitsuhashi, W., et al. (2004). Biological functions of ent- and syn-copalyl diphosphate synthases in rice: key enzymes for the branch point of gibberellin and phytoalexin biosynthesis. *Plant J.* 39, 886–893. doi: 10.1111/j.1365-313X.2004.02175.x
- Phukan, U. J., Jeena, G. S., and Shukla, R. K. (2016). WRKY transcription factors: molecular regulation and stress responses in plants. *Front. Plant Sci.* 7:760. doi: 10.3389/fpls.2016.00760
- Qu, R., Cao, Y., Tang, X., Sun, L., Wei, L., and Wang, K. (2021). Identification and expression analysis of the WRKY gene family in *Isatis indigotica*. *Gene* 783:145561. doi: 10.1016/j.gene.2021.145561
- Rinerson, C. I., Rabara, R. C., Tripathi, P., Shen, Q. J., and Rushton, P. J. (2015). The evolution of WRKY transcription factors. *BMC Plant Biol.* 15:66. doi: 10.1186/s12870-015-0456-y
- Rodrigues, F. A., McNally, D. J., Datnoff, L. E., Jones, J. B., Labbé, C., Benhamou, N., et al. (2004). Silicon enhances the accumulation of diterpenoid phytoalexins in rice: a potential mechanism for blast resistance. *Phytopathology* 94, 177–183. doi: 10.1094/PHYTO.2004.94.2.177
- Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY transcription factors. *Trends Plant Sci.* 15, 247–258. doi: 10.1016/j.tplants.2010.02.006
- Ryu, H. S., Han, M., Lee, S. K., Cho, J. I., Ryoo, N., Heu, S., et al. (2006). A comprehensive expression analysis of the WRKY gene superfamily in rice plants during defense response. *Plant Cell Rep.* 25, 836–847. doi: 10.1007/s00299-006-0138-1
- Tamogami, S., and Kodama, O. (2000). Coronatine elicits phytoalexin production in rice leaves (*Oryza sativa* L.) in the same manner as jasmonic acid. *Phytochemistry* 54, 689–694. doi: 10.1016/s0031-9422(00)00190-4
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Toyomasu, T., Usui, M., Sugawara, C., Otomo, K., Hirose, Y., Miyao, A., et al. (2014). Reverse-genetic approach to verify physiological roles of rice phytoalexins: characterization of a knockdown mutant of OsCPS4 phytoalexin biosynthetic gene in rice. *Physiol. Plant.* 150, 55–62. doi: 10.1111/ppl.12066
- Wang, Q., Hillwig, M. L., and Peters, R. J. (2011). CYP99A3: functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice. *Plant J.* 65, 87–95. doi: 10.1111/j.1365-313X.2010.04408.x
- Wang, Y., Zhang, Y., Fan, C., Wei, Y., Meng, J., Li, Z., et al. (2021). Genome-wide analysis of MYB transcription factors and their responses to salt stress in *Casuarina equisetifolia*. *BMC Plant Biol.* 21:328. doi: 10.1186/s12870-021-03083-6
- Wilderman, P. R., Xu, M., Jin, Y., Coates, R. M., and Peters, R. J. (2004). Identification of syn-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. *Plant Physiol.* 135, 2098–2105. doi: 10.1104/pp.104.045971
- Yamada, Y., Nishida, S., Shitan, N., and Sato, F. (2021). Genome-wide profiling of WRKY genes involved in benzylisoquinoline alkaloid biosynthesis in California Poppy (*Eschscholzia californica*). *Front. Plant Sci.* 12:699326. doi: 10.3389/fpls.2021.699326
- Yokotani, N., Sato, Y., Tanabe, S., Chujo, T., Shimizu, T., Okada, K., et al. (2013). WRKY76 is a rice transcriptional repressor playing opposite roles in blast disease resistance and cold stress tolerance. *J. Exp. Bot.* 64, 5085–5097. doi: 10.1093/jxb/ert298
- Zhang, L., Chen, C., Xie, F., Hua, Q., Zhang, Z., Zhang, R., et al. (2021). A novel WRKY transcription factor HmoWRKY40 associated with betalain biosynthesis in Pitaya (*Hylocereus monacanthus*) through regulating HmoCYP76AD1. *Int. J. Mol. Sci.* 22:2171. doi: 10.3390/ijms22042171
- Zhang, R., Chen, Z., Zhang, L., Yao, W., Xu, Z., Liao, B., et al. (2021). Genomic characterization of WRKY transcription factors related to andrographolide biosynthesis in *Andrographis paniculata*. *Front. Genet.* 11:601689. doi: 10.3389/fgene.2020.601689
- Zhang, Y., and Wang, L. (2005). The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol. Biol.* 5:1. doi: 10.1186/1471-2148-5-1
- Zhao, X. K., Fan, X. Q., Nie, M. E., Wang, H. Y., Liu, Q. S., and Duan, Y. H. (2020). Identification and bioinformatics analysis of WRKY family in *Sorghum*. *Mol. Plant Breed.* 13, 4170–4181. doi: 10.13271/j.mpb.018.004170
- Zhou, Q. Y., Tian, A. G., Zou, H. F., Xie, Z. M., Lei, G., Huang, J., et al. (2008). Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. *Plant Biotechnol. J.* 6, 486–503. doi: 10.1111/j.1467-7652.2008.00336.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Zhu, Shi, Huang, Zhang, Kai and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.