

Article

Bioinformatics Analysis of WRKY Family Genes in *Erianthus fulvus* Ness

Haowen Chen ^{1,†}, Xuzhen Li ^{2,3,†}, Fusheng Li ¹, Dengyu Li ¹, Yang Dong ^{2,3,*} and Yuanhong Fan ^{4,*}¹ College of Agronomy and Biotechnology, Yunnan Agricultural University, Kunming 650201, China² State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming 650201, China³ College of Biological Big Data, Yunnan Agricultural University, Kunming 650201, China⁴ Yunnan Plateau Characteristic Agricultural Industry Research Institute, Yunnan Agricultural University, Kunming 650201, China

* Correspondence: loyalyang@163.com (Y.D.); 2017666@ynau.edu.cn (Y.F.)

† These authors contributed equally to this work.

Abstract: One of the most prominent transcription factors in higher plants, the WRKY gene family, is crucial for secondary metabolism, phytohormone signaling, plant defense responses, and plant responses to abiotic stresses. It can control the expression of a wide range of target genes by coordinating with other DNA-binding or non-DNA-binding interacting proteins. In this study, we performed a genome-wide analysis of the *EfWRKY* genes and initially identified 89 members of the *EfWRKY* transcription factor family. Using some members of the *OsWRKY* transcription factor family, an evolutionary tree was built using the neighbor-joining (NJ) method to classify the 89 members of the *EfWRKY* transcription factor family into three major taxa and one unclassified group. Molecular weights ranged from 22,614.82 to 303,622.06 Da; hydrophilicity ranged from (−0.983)–(0.159); instability coefficients ranged from 40.97–81.30; lipid coefficients ranged from 38.54–91.89; amino acid numbers ranged from 213–2738 bp; isoelectric points ranged from 4.85–10.06. A signal peptide was present in *EfWRKY41* but not in the other proteins, and *EfWRK85* was subcellularly localized to the cell membrane. Chromosome localization revealed that the WRKY gene was present on each chromosome, proving that the conserved pattern WRKYGQK is the family's central conserved motif. Conserved motif analysis showed that practically all members have this motif. Analysis of the cis-acting elements indicated that, in addition to the fundamental TATA-box, CAAT-box, and light-responsive features (GT1-box), there are response elements implicated in numerous hormones, growth regulation, secondary metabolism, and abiotic stressors. These results inform further studies on the function of *EfWRKY* genes and will lead to the improvement of sugarcane.

Keywords: *Erianthus fulvus* ness; WRKY gene family; structural domains; bioinformatics analysis

Citation: Chen, H.; Li, X.; Li, F.; Li, D.; Dong, Y.; Fan, Y. Bioinformatics Analysis of WRKY Family Genes in *Erianthus fulvus* Ness. *Genes* **2022**, *13*, 2102. <https://doi.org/10.3390/genes13112102>

Academic Editors: M. Michael Gromiha and Y-h. Taguchi

Received: 13 October 2022

Accepted: 7 November 2022

Published: 12 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sugarcane is a C4 plant of the grass family [1]. It is used as a raw material to manufacture sucrose and to refine ethanol for biomass energy and livestock feed [2]. Sugarcane has a 12 to 18-month reproductive life, and seasonal and lengthy droughts are unavoidable. Water deficiency directly affects the tillering and elongation of sugarcane at the seedling stage, resulting in lower yields [3]. In China, 80 percent of sugarcane is grown in semi-arid or dry areas where droughts occur frequently and cause significant losses [4–6]. As a result, it is critical to thoroughly investigate sugarcane's drought resistance mechanism, tap its drought resistance genes, and produce exceptional drought-tolerant types. *E. fulvus* Ness is a wild species of Cane grass, a near relative of sugarcane, which grows on high-altitude slopes in arid, barren, and cold environments, even in the cracks of steep rock walls. Therefore, it is exceptionally tolerant of drought, barrenness, and cold compared to other sugarcane varieties [7].

WRKY transcription factors constitute one of the most prominent families of transcriptional regulators in plants and are involved in regulating various plant functions through signaling networks [8]. The name of the *WRKY* family comes from the *WRKY* domain, a region consisting of 60 amino acids highly conserved among family members, which is the most noticeable feature of their proteins [9]. The conserved amino acid sequence *WRKYGQ(E/K)K* and a zinc finger motif at the N terminus of the *WRKY* structural domain serve as its defining features [10]. Since all characterized *WRKY* proteins bind the identical DNA motif, it is assumed that the *WRKY* structural domain functions as the DNA binding domain's single conserved structural component [11]. Every one of the known *WRKY* proteins has one or two *WRKY* structural domains. The first classification of *WRKY* proteins was into two major groups [12], with Group I consisting of proteins with two *WRKY* structural domains and Group II consisting of proteins with one *WRKY* structural domain. Members of Groups I and II typically have the same zinc finger structure [13]. Only a tiny part of the *WRKY* proteins contain zinc finger architectures distinct from Groups I and II. As a result, they are categorized in the newly established Group III. However, it has been demonstrated that individuals from all three groups selectively bind to different W-box elements [14].

Ishiguro and Nakamura identified the first *WRKY* transcription factor, *SPF1*, from sweet potatoes [15]. In recent years, *WRKY* transcription factors have been increasingly analyzed in other plants, such as cucumber [16], tomato [17], and chickpea [18]. As *WRKY* transcription factors have been studied in depth, some of their roles in plants have been gradually explored. *NaWRKY3* in tobacco is required to induce the *NaWRKY6* gene by fatty acid-amino acid conjugates in the oral secretions of the tobacco moth (*Manduca sexta*) larvae, and silencing one or both of these genes makes plants highly susceptible to herbivore attack. Both genes appear to help plants distinguish between mechanical injury and herbivore attack [19]; *Arabidopsis* plants with *AtWRKY23* taken out exhibit excellent resistance to encapsulated nematodes because it is up-regulated right away after nematode infection [20]. The *WRKY* gene, isolated from the dry evergreen C3 plant creosote shrub, was first demonstrated to abscisic acid (ABA) signaling in one of the earliest studies [21]. Heat and drought tolerance were improved by heat stress-induced overexpression of *OsWRKY11* mediated by the *HSP101* promoter [21]. And similarly, overexpression of *OsWRKY45* enhanced salt and drought tolerance in addition to disease resistance [22]. Compared to wild-type plants, *Arabidopsis* plants overexpressing *GmWRKY21* in soybean were more resistant to cold stress, and *GmWRKY54* in soybean was more resistant to salt and drought. Additionally, *WRKY* genes play roles in seed germination and dormancy [23], seed development [24,25], and anti-aging [26,27].

In this study, we comprehensively analyzed the physicochemical properties, secondary structure, conserved motifs, gene structure, cis-acting elements, gene duplication events, and GO enrichment analysis of the *WRKY* gene family in *E. fulvus*. Phylogenetic analysis of the evolutionary relationship of the *WRKY* gene family between *E. fulvus* and rice was performed. Finally, we analyzed the homology of *WRKY* genes in *Arabidopsis thaliana*, *Sorghum bicolor*, *Miscanthus iturioriparius*, *Oryza sativa*, and *Saccharum spontaneum*. This study will provide valuable information for further studies on the function of *WRKY* genes in *E. fulvus* and subsequent studies on molecular mechanisms.

2. Materials and Methods

2.1. Identification of the *EfWRKY* Gene Family Members

Genomic data, protein files, CDS files, and gff3 files of *E. fulvus* were downloaded from the EfGD: *E. fulvus* Genome Database (<https://efgenome.ynau.edu.cn>) (accessed on 25 August 2022) online website [28]. The genomes of *Arabidopsis thaliana* (GenBank: GCA_000001735.2), *Sorghum bicolor* (GenBank: GCA_000003195.3), *Miscanthus iturioriparius* (GenBank: GCA_904845875.1), *Oryza sativa* (GenBank: GCA_001433935.1), and *Saccharum spontaneum* (GenBank: GCA_022457205.1) with gff3 files were downloaded from NCBI. The HMM profile of the *WRKY* conserved structural domain (PF03106) was searched

and downloaded from the Pfam database (<http://pfam.xfam.org/>) (accessed on 3 May 2022). The search command in HMMER 3.0 [29] software and the HMM profile of the WRKY conserved structural domain were used to search the sucralose protein files (E-value 10^{-5}). A total of 90 WRKY genes were searched, and the sequences of the 90 genes in *E. fulvus* were submitted to Batch CD-Search on NCBI (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) (accessed on 26 August 2022) and to the SMART online website (http://smart.embl-heidelberg.de/smart/set_mode) (accessed on 26 August 2022) for structural domain analysis, and genes that did not contain WRKY structural domains were removed, yielding a total of 89 WRKY genes. They were renamed *EfWRKY01–EfWRKY89*.

2.2. Predicted Physicochemical Properties and Secondary Structure Analysis of WRKY Protein in *E. fulvus*

Prediction of amino acid number (AA), isoelectric point (PI), molecular weight (MW), and hydrophilicity (GRAVY) of WRKY proteins in *E. fulvus* were conducted using the online analysis website (<https://web.expasy.org/protparam/>) (accessed on 27 August 2022); the Plant-mPLoc server (<http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/>) (accessed on 28 August 2022) was used for subcellular localization prediction; SignalP-5.0 (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>) (accessed on 28 August 2022) for signal peptide prediction (Parameter setting: organism group is eukaryotic); SOPMA (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) (accessed on 29 August 2022) for analysis of the secondary structure of WRKY proteins in *E. fulvus* (Parameter setting: the number of conception states is 4, and the similarity threshold is 8).

2.3. Construction of a Phylogenetic Tree

The WRKY protein sequences of *EfWRKY* genes were subjected to multiple sequence alignment using Jalview software (Mafft with Defaults) and, after construction, were grouped according to the number of WRKY structural domains, the type of zinc finger structures. To clarify the evolutionary relationships between individual members of the *EfWRKY* gene family, 89 *EfWRKYs* genes were compared with 83 *OsWRKYs* genes using the Jalview software (Mafft with Defaults). The results were saved, and a phylogenetic tree was constructed using the neighbor-joining method (NJ) in MEGA 7.0 [30] (Parameter settings: Bootstrap method 1000; p-distance; uniform rates; Pairwise deletion).

2.4. Analysis of Conserved Motifs, Gene Structure, and Chromosomal Localization of WRKY Proteins in *E. fulvus*

Conserved motif prediction for 89 WRKY proteins in *E. fulvus* used MEME [31] (<https://meme-suite.org/meme/>) (accessed on 27 August 2022); Parameter settings: allowing for recurrence of conserved structural domains and setting the number of conserved motifs to 10; mapping of conserved motifs, gene structure, and gene chromosome positioning of WRKY proteins in *E. fulvus* used TBtools software (for version 0.665).

2.5. Chromosome Localization, Gene Duplication, and Covariance Analysis

Two gene amplification mechanisms—tandem duplication and segmental duplication—were considered to examine gene duplication occurrences. The *E. fulvus* genome files were swiftly compared with the gff3 files using the TBtools software (Parameter setting: CPU for BlastP is 2, E-value 10^{-5} , Num of BlastHits is 10). Tandem duplications were defined as two gene pairs in the same chromosomal region separated by five or fewer loci. Used TBtools software, Ka (heteronymous substitutions) and Ks (synonymous substitutions) were estimated to evaluate the impact of selection pressure and dispersal time on *EfWRKY* genes (Parameter setting: CPU is 4). To assess the conservation of segmental duplicate gene pairs and direct homologs of *EfWRKY* with other plant species, TBtools software generated visual covariance maps.

2.6. Analysis of Cis-Acting Elements of Promoters

A 2000 bp upstream of the gene CDS sequence was extracted and submitted to the PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (accessed on 27 August 2022) website for cis-acting element prediction, with some unclear elements removed, using TBtools [32] software to draw.

2.7. Gene Ontology (GO) Annotation

The Swiss-Prot files were downloaded from the Uniprot website and compared used TBtools software to obtain GO data for 89 *EfWRKY* genes; the GO annotation files for all genes in sucrose were obtained from the EfDG website as background files, and the gene ontology annotations for the *EfWRKY* genes were analyzed using the online website (<https://www.omicshare.com/tools/Home/Soft/gogseasenor>) (accessed on 10 September 2022).

3. Results

3.1. Identification and Prediction of Physicochemical Properties of *EfWRKY* Gene Family Members

89 *WRKY* transcription factors were identified in *E. fulvus* (Table S1). The 89 *WRKY* proteins in *E. fulvus* ranged in amino acid length from 213 to 2738 amino acids (aa), molecular weight from 22,614.82 to 303,622.06, the isoelectric point from 4.85 to 10.06, instability coefficient from 40.97 to 81.30, and lipid coefficient from 38.54 to 91.98; all other sub-cells, except for *EfWRKY85*, were localized on the nucleus; all the rest were without signal peptide except *EfWRKY41*, which had a signal peptide (Table S2).

3.2. Secondary Structure Analysis of the 89 *WRKY* Protein in *E. fulvus*

The study of the secondary structure of the 89 *WRKY* proteins in *E. fulvus* revealed that its secondary structure mainly consisted of irregular coils, α -helices, β -turns, and extended chains, which made it easier to understand the spatial structure of the protein through data analysis. Among these four secondary structure elements, the irregularly coiled coil accounted for the highest percentage, and its function was mainly to connect other secondary structure elements, which ranged from 35.83% (*EfWRKY85*) to 76.67% (*EfWRKY09*). In contrast, the β -turn angle accounted for a smaller percentage of these four structures. The secondary structures of the sucralose *WRKY* family are relatively neat, except for *EfWRKY31*, *EfWRKY61*, and *EfWRKY01*, which are irregularly curled > extended chain > α -helix > β -turn. In contrast, all other family members are irregularly curled > α -helix > extended chain > β -turn (Table S3).

3.3. Phylogenetic, Conserved Motif, and Gene Structure Analysis of *WRKY* Gene Family Members in *E. fulvus*

The 89 *WRKY* protein sequences of *EfWRKY* genes can be divided into three major groups of five subgroups and one non-class group (Figure 1). The proteins in Group I contain two *WRKY* conserved structural domains (*WRKYGQK*/*WRKYGEK*) and zinc finger structures *CX4CX22HX1H* or *CX4CX23HX1H*; the proteins in Group II have one *WRKY* conserved structural domain (*WRKYGQK*/*WRKYGKK*/*WEKFGK*) and a zinc finger structure *CX5CX23HX1H*. The proteins in Group II were further divided into five subgroups; the proteins in Group III contained one *WRKY* conserved structural domain (*WRKYGQK*/*WRKYGEK*) and the zinc finger structure *CX5-8CX23HX1C*; the remaining three belonged to Group NG. In the evolutionary tree analysis of the *WRKY* and rice *WRKY* gene families in *E. fulvus*, it was found that Group II occupied the majority (50.56%), followed by Group III (30.34%). The minor proportion was Group I (14.61%); the first *WRKY* structural domain K of the *EfWRKY07* gene was mutated to *WRKYGEK*, and the same mutations occurred in *EfWRKY31*, *EfWRKY45*, *EfWRKY50*, *EfWRKY54*, and *EfWRKY83*; the conserved *WRKY* structure of *EfWRKY40*, *EfWRKY53*, *EfWRKY57*, and *EfWRKY67* was mutated to *WRKYGKK* (Figure 1). This phenomenon is consistent with the variation in *Dendrobium officinale* [33].

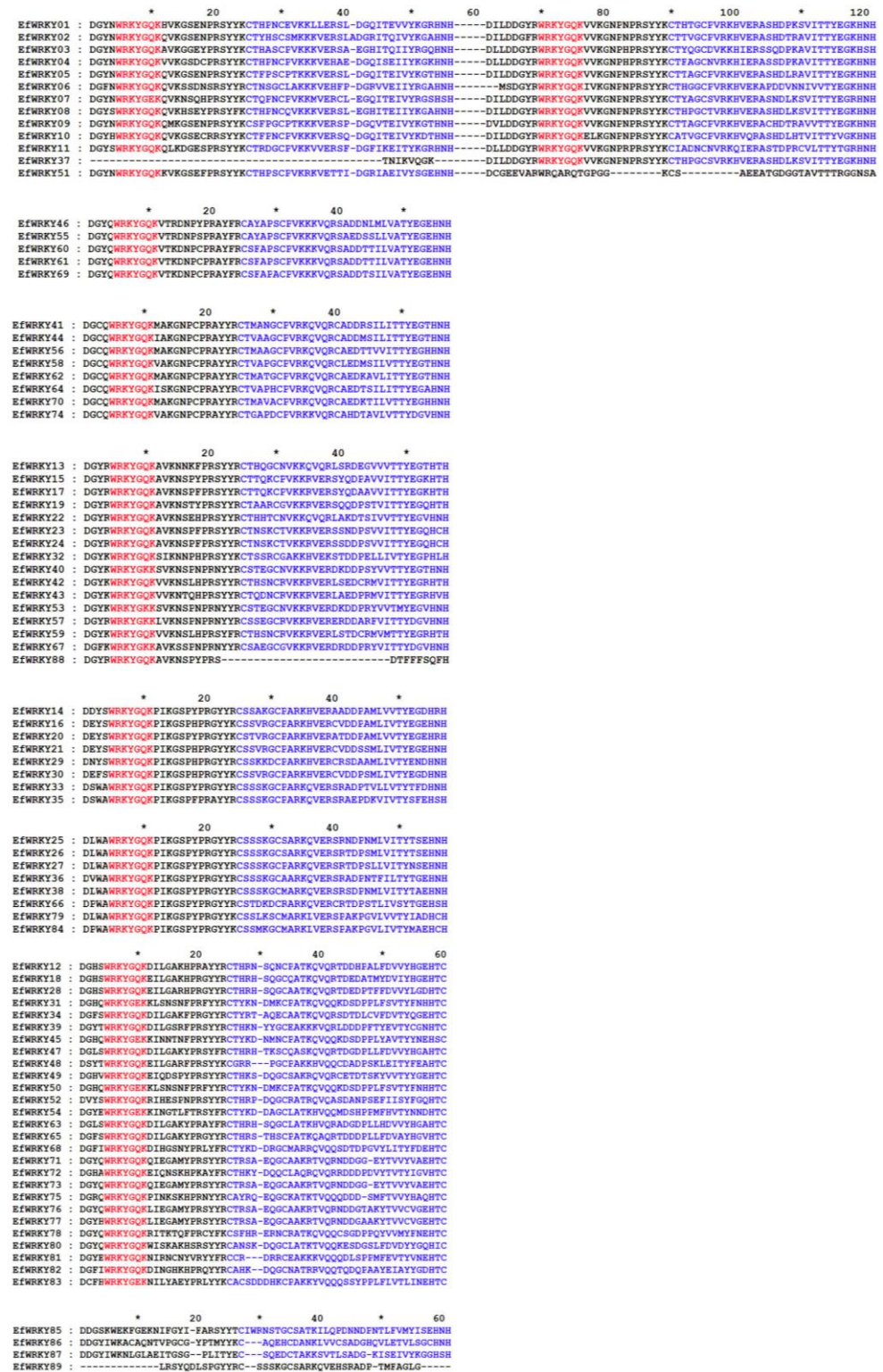


Figure 1. Sequence alignment of the WRKY structural domain of the 89 WRKY proteins in *E. fulvus*. The red part is the highly conserved heptapeptide of the protein, and the blue part is the zinc finger structure of the protein.

Conserved motifs were predicted for 89 WRKY proteins in *E. fulvus* using MEME [31] (<https://meme-suite.org/meme/>) (accessed on 27 August 2022), allowing for the recurrence of conserved structural domains and setting the number of conserved motifs to 10; distribution of conserved motifs and gene structure analysis revealed that the motif1 motif

was highly conserved; two motifs 7 were present in *EfWRKY89*, two motifs 4 in *EfWRKY11*, and only one conserved motif 6 in *EfWRKY87* (Table S4). The phylogenetic tree among the members of the *EfWRKY* gene family is shown in Figure 2A. There are differences in the conserved motifs between members of the *EfWRKY* gene family, which may also make the functions of the gene family members differ from each other (Figure 2B); Gene structure analysis revealed that all members of the *EfWRKYs* family had introns ranging from 1–9 in number, except for *EfWRKY12*, *EfWRKY14*, and *EfWRKY54*, which had no introns, and some members of the *EfWRKY* gene family did not contain untranslated regions (UTRs) at the 5' end and 3' end (Figure 2C).

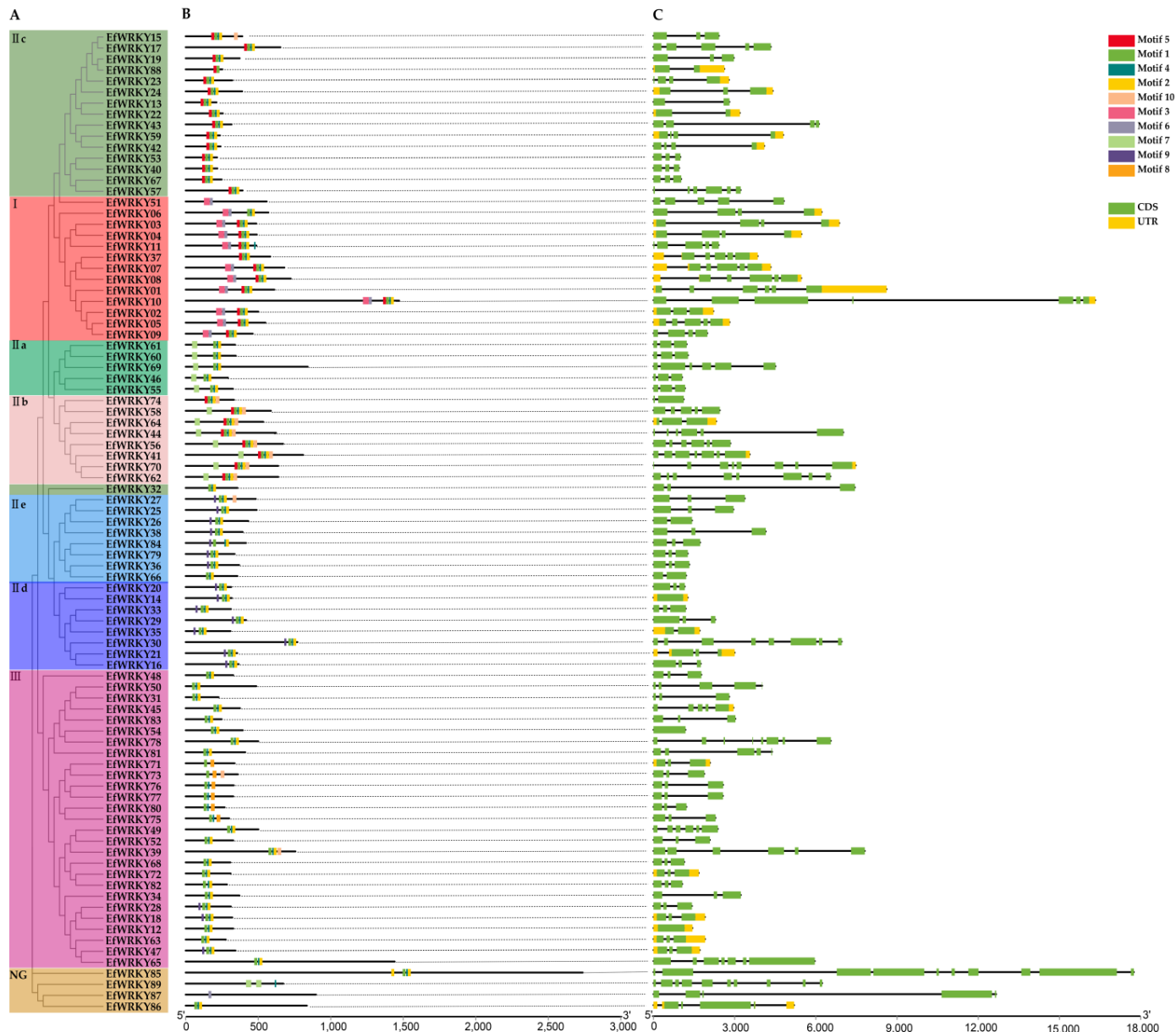


Figure 2. Conserved motifs and gene structure maps of 89 *EfWRKY* genes. (A) Phylogenetic tree of 89 *EfWRKY* genes. (B) Motif composition of 89 *EfWRKY* genes; a total of 10 types of conserved domains were found. The black line indicates the protein length. (C) Gene structure of 89 *EfWRKY* genes. Black lines indicate introns, green boxes indicate CDS, and yellow boxes indicate UTR.

3.4. Phylogenetic and Functional Analysis of the OsWRKY and EfWRKY Gene Families

To further investigate the different functions of the *EfWRKY* genes, the sequences of the WRKY proteins in *E. fulvus* and *O. sativa* were aligned (Table S5), and a phylogenetic tree was constructed using the NJ (Neighbor-Joining) (Figure 3). Of the 83 *O. sativa* WRKY proteins, there are 11 with known functions. As can be seen from Figure 3, members of Group I of the *EfWRKY* gene family cluster with *OsWRKY89* due to high protein homology, with the *EfWRKY06* and *OsWRKY89* genes being closely related evolutionarily and likely to play an essential role in biotic and abiotic stress responses [34]; members of *EfWRKY* gene family group IIa clustered with *OsWRKY76* and *OsWRKY71*, and it is possible that members of group IIa in the *EfWRKY* gene family play an active role in plant cold tolerance [35]; members of Group IIb of the *EfWRKY* gene family clustered with *OsWRKY5*, with a strong evolutionary relationship between *EfWRKY62* and the *OsWRKY5* gene, which may have a role in plant leaf senescence [36]; members of Group IIc of the *EfWRKY* gene family cluster with *OsWRKY11*, *OsWRKY29* and *OsWRKY72*, with *EfWRKY13* and *OsWRKY72* genes, being closely related evolutionarily, and may cause endogenous plant jasmonic acid (JA) levels in plants, ultimately causing susceptibility to white leaf blight; there is a close evolutionary relationship between *EfWRKY24* and *OsWRKY29* genes, which may block abscisic acid signaling and inhibit seed dormancy in rice [37]; and there is a close evolutionary relationship between *EfWRKY88* and *OsWRKY11* genes, which may play an essential role in heat and drought stress response and resistance [38]. Members of Group IId of the *EfWRKY* gene family cluster with *OsWRKY13*, with a strong evolutionary relationship between *EfWRKY33* and *OsWRKY13* genes that may have a function in disease resistance in plants [39]; members of Group III of the *EfWRKY* gene family cluster with *OsWRKY47* and *OsWRKY74* are clustered together, with a strong evolutionary relationship between *EfWRKY52* and *OsWRKY47* genes, which may be a positive regulator of water deficit stress response [40], and *EfWRKY18* and *OsWRKY74* genes, which may have a role in plant response to low temperature or in response to iron (Fe) and nitrogen (N) deficiency [40].

3.5. Analysis of Cis-Acting Elements of 89 *EfWRKY* Gene Promoters

In addition to the basic TATA-box, CAAT-box, and light-responsive features (GT1-box), there are also response elements involved in various hormones, growth regulation, secondary metabolism, and abiotic stresses (Tables S6 and S7) (Figure 4B).

3.6. Chromosomal Localization Analysis of 89 *EfWRKY* Genes

Chromosomal localization analysis revealed that 89 *EfWRKY* genes were unevenly distributed on ten chromosomes, with the most extensive distribution of 20 WRKY genes on Chr4 and the minor distribution of 5 genes on Chr5, 6, and 9; according to Holub [41] et al., chromosomal regions containing two or more genes within 200 kb are defined as tandem duplication events. Tandem and segmental repeats contribute to the expansion of new gene family members and new functions in the evolution of plant genomes. To study gene duplication events in the *E. fulvus* genome, we investigated tandem and segmental duplication during the transition of the *EfWRKY* gene family. A total of 2 *EfWRKY* gene pairs were confirmed to be tandem duplicates. Based on gene duplication analysis, chr1 and chr7 were observed to undergo a tandem duplication gene pair; in addition to the tandem duplication, there were 22 segmental duplications (Figure 5). To determine the selective evolutionary pressure of *EfWRKY* gene divergence after repetition, the K_a and K_s values of the duplicated *EfWRKY* gene pairs were calculated using the KaKs calculator. In general, $K_a/K_s = 1$ indicates neutral selection, $K_a/K_s > 1$ indicates positive selection, and $K_a/K_s < 1$ indicates purifying selection. $K_a/K_s < 1$ for each duplicated *EfWRKY* gene pair means purifying selection during evolution (Table S8).

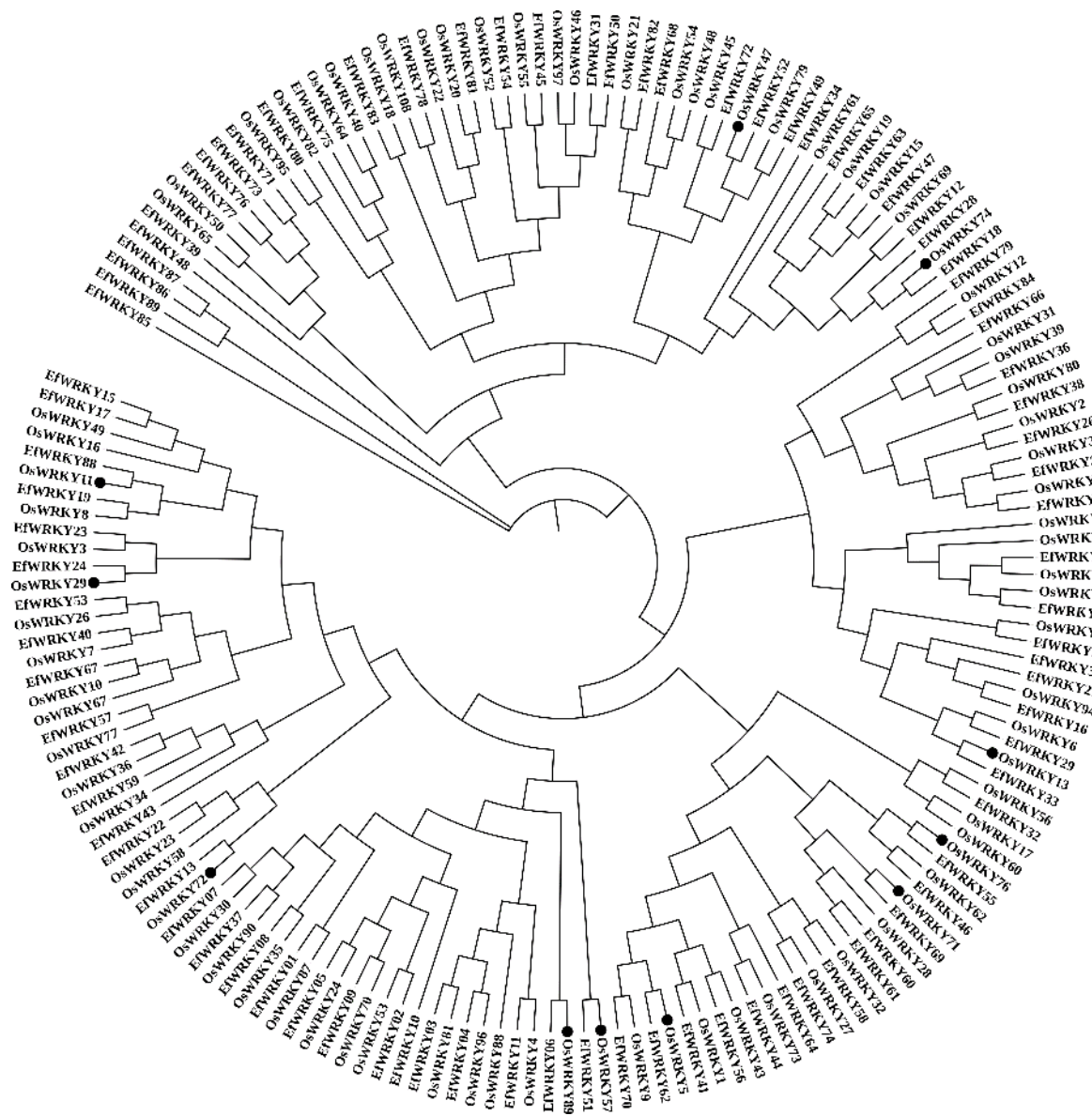


Figure 3. Phylogenetic tree of WRKY protein sequences in *E. fulvus* and some *O. sativa*. (● indicates known WRKY genes in rice).

3.7. Synteny Relationship of *EfWRKY* Genes

To further infer evolutionary relationships, WRKY genes were compared to identify direct homologous *EfWRKY* gene pairs among *E. fulvus*, *A. thaliana*, *S. bicolor*, *M. iturioriparius*, *O. sativa*, and *S. officinarum*. Based on the homozygosity analysis, 15 of the 89 *EfWRKY* genes had common lineage pairs in Arabidopsis; 66 had corresponding direct homologs in *O. sativa* and *S. spontaneum*; 70 had related direct homologs in *S. bicolor*, and 71 had corresponding direct homologs in *M. iturioriparius* (Table S9). The co-linear plot shows that *EfWRKY* genes are highly evolutionarily homologous to sorghum and manzanita WRKY genes, implying that they may have related functions (Figure 6).



Figure 4. Cis-element analysis of the *EfWRKY* gene family. Different colors represent a total of 26 types of functional modules. **(A)** Phylogenetic tree of 89 *EfWRKY* genes. Numbers on branches represent bootstrap estimates for 1000 replicate analysis or clade credibility values. **(B)** Cis-acting elements.

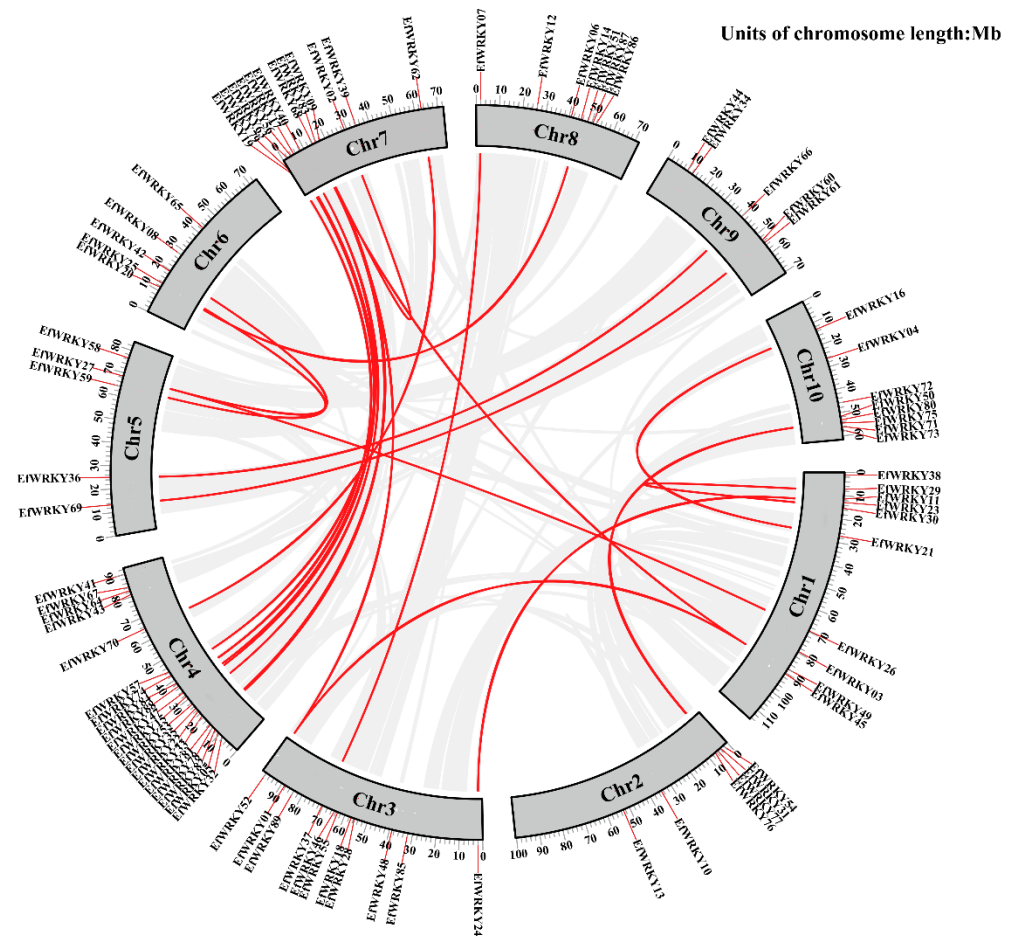


Figure 5. The intrachromosomal segmental duplication map of the *WRKY* genes in *E. fulvus*. Red lines are *WRKY* gene pairs genomes between plant genomes. Gray lines are collinear blocks for interplant genomes.

3.8. Gene Ontology (GO) Annotation

To explore the function of the *EfWRKY* gene in different biological processes, molecular procedures, and cell compartment construction, GO data for 89 *E. fulvus* (Table S10) were submitted as background files with the EfGD: *E. fulvus* Genome Database (<https://efgenome.yнау.edu.cn/index/index/index>) (25 August 2022) of all *E. fulvus* gene GO data (Table S11) as background files submitted online at (<https://www.omicshare.com/tools/Home/Soft/gogseasenor>) (accessed on 10 September 2022). GO enrichment analysis was performed. As shown in Figure 7, GO function annotation was performed, and, as provided in data s8, gene ontology numbers were also identified during the analysis. Biological processes suggest that *EfWRKY* genes are involved in transcriptional regulation, regulating processes such as biosynthesis and gene expression. Most *EfWRKY* genes are involved in sequence-specific DNA binding functions and transcriptional regulation. It is inferred that they are essential as transcription factors for the growth and development of *E. fulvus* and for regulating the expression of stress-responsive genes.

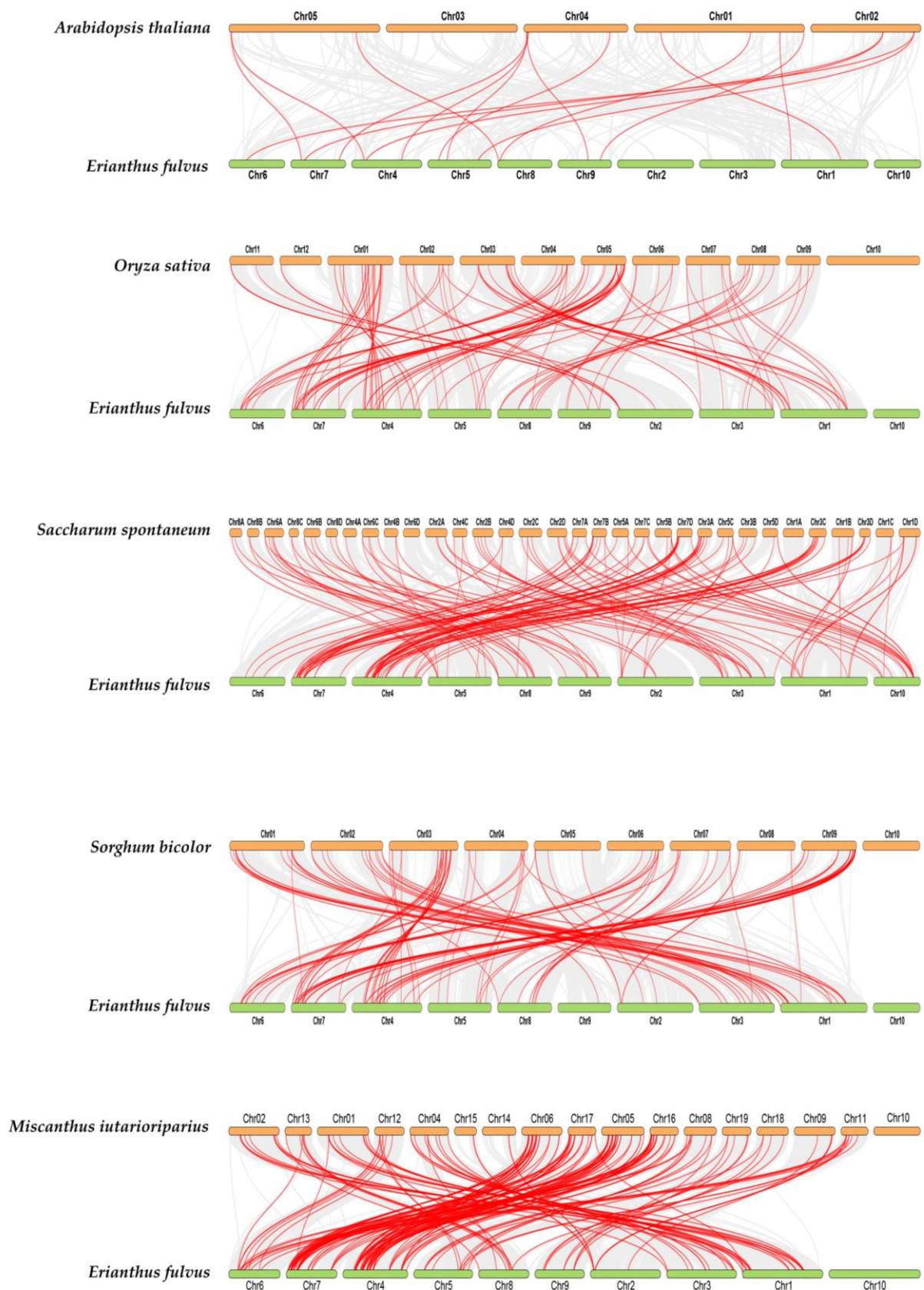


Figure 6. Synteny analysis of WRKY family. *A. thaliana*, *S. bicolor*, *M. itarioriparius*, *E. fulvus*, *O. sativa*, and *S. spontaneum*. Red lines are the syntenic WRKY gene pairs. Gray lines are collinear blocks.

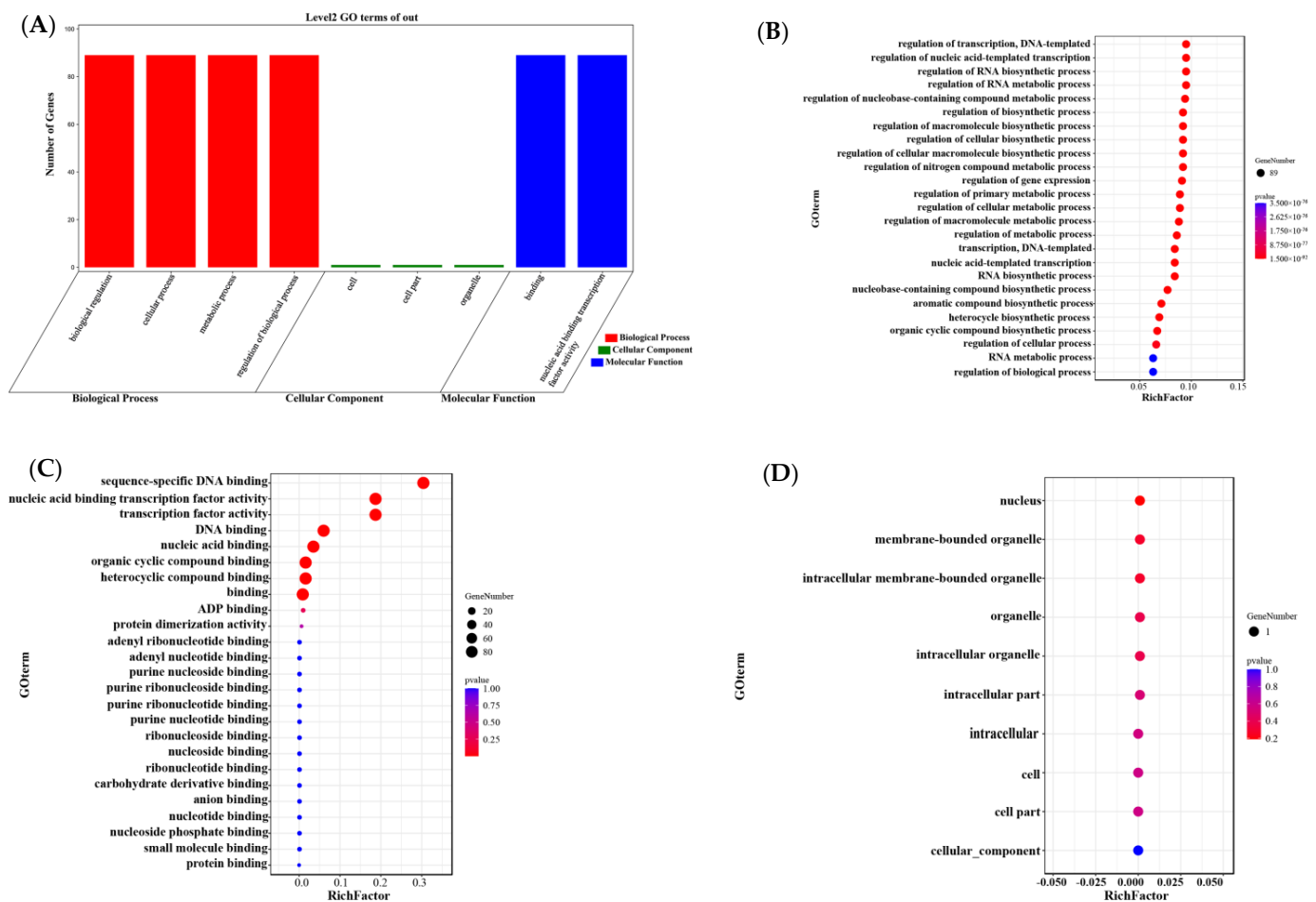


Figure 7. Gene Ontology (GO) annotation. (A) Direct count GO. Red bars represent biological processes (BP), blue bars represent molecular functions (MF) and green bars represent cellular compartments (CC); (B) Major Biological Process (BP); (C) Major Molecular Function (MF); (D) Cell Compartment (CC).

4. Discussion

Crop improvement requires the functional, genomic and structural characterization of individual gene families involved in regulating plant metabolism and developmental processes [42,43]. Although many methods have been developed to describe gene family members, these methods often require considerable effort [44–46]. Identifying related gene families and performing careful annotation work is necessary to obtain genes with potentially specialized functions. Many plant genomes have been sequenced, which can be aided by data mining and phylogenetic techniques. In recent years, research has focused on genome-wide analysis of plant gene families, particularly in crops [47–50], and *WRKY* is one of the larger gene families in plants whose role is regulating plant growth and development [51–53]. Similarly, analysis of biotic and abiotic stresses has been demonstrated [54–56].

In this study, we identified a total of 89 members of the *EfWRKY* gene family using the pep file of the *E. fulvus* genome and the HMM profile of the conserved structural domain of *WRKY* (PF03106), with individual members having almost identical gene structures (Figure 2), the *WRKYGQK* heptapeptide structure with the zinc finger structure (C2H2 type or C2HC type). The difference lies in the number of *WRKY* domains. As seen from Figure 2, *EfWRKY37* and *EfWRKY51* have a *WRKY* domain but are assigned to Group I, probably missing during evolution, which usually occurs in monocotyledons [57].

Meanwhile, the WRKYGQK heptapeptide structures of *EfWRKY40*, *EfWRKY53*, *EfWRKY57*, and *EfWRKY67* mutated to WRKYGKK, a variant also found in rice and dates; and the WRKYGQK heptapeptide structure of *EfWRKY31*, *EfWRKY45*, *EfWRKY50*, *EfWRKY54*, and *EfWRKY83* mutated to WRKYGEK. It has been demonstrated that the WRKYGKK structural domain binds specifically to the WK-box, which is significantly different from the WRKYGQK-specific binding [58]. Variants in conserved WRKY sequences may affect the regular interaction and binding specificity of WRKY genes with downstream target genes, thereby altering gene function [59].

Phylogenetic analysis revealed that the members of the *EfWRKY* gene family conformed to the consistent classification of WRKY gene families, except for a few *E. fulvovus* WRKY proteins with partially missing or mutated structural domains. All other members had complete and accurate structural parts, indicating to some extent that the *EfWRKY* gene family is relatively evolutionarily conserved. As can be seen from Figure 1, the *EfWRKY* gene family is divided into four groups: Group I, Group II, Group III, and NG; the number of its groups containing WRKY genes is 13 (14.61%), 45 (50.56%), 27 (30.34%) and 4 (4.49%), respectively. Group II is the group with the highest number of *EfWRKY* genes, indicating that the smallest number of Group IIa WRKY genes may be due to its relatively short evolutionary time compared to the other groups, being the last of all groups in the *EfWRKY* gene family to evolve. This explains the variation and deletion of conserved structural domains of WRKY proteins in *E. fulvovus*, suggesting that *E. fulvovus* may be gaining evolution.

In conclusion, systematic analysis of the *EfWRKY* gene family could provide a valuable reference for more detailed functional genomic characterization of these genes and help to predict and select suitable candidate genes for sugarcane improvement. As more complete sequenced genomes become available, more WRKY genes will likely be identified in other plants. Based on these data, a more comprehensive understanding of the origin and evolution of the WRKY gene family will probably be found in the data identified, and its functional characteristics can be obtained.

5. Conclusions

In conclusion, we performed a genome-wide identification of WRKY gene families in cane grass. Using bioinformatics tools, we investigated their evolutionary relationships, gene structure, replication events, and cis-acting promoter progenitors. Eighty-nine WRKY genes with WRKY structural domains were identified in the *E. fulvovus* genome and distributed unevenly on ten chromosomes. Based on sequence alignment and phylogenetic analysis, *EfWRKY* genes were classified into three prominent families, five subgroups, and one non-main family. The gene structure and conserved motif analysis were also consistent with the phylogenetic classification. The gene structure and motif patterns indicated that *EfWRKY* members in the same subgroup showed extensive similarity. Tandem and segmental duplication significantly contributed to the expansion of the WRKY gene family in *E. fulvovus*. Evolutionary divergence analysis (Ka/Ks) indicated that *EfWRKYs* were subject to strong purifying selection during plant evolution. Homology analysis showed that 70 and 71 *EfWRKY* genes were directly homologous to *S. bicolor* and *M. iturioriparius*, respectively. In addition, subcellular localization revealed that 88 *EfWRKYs* genes were mainly located in the nucleus. Overall, our findings describe the evolutionary characteristics, genomic replication events, and molecular biological functions of WRKY genes in cane grass. In conclusion, our results will provide a basis for elucidating the molecular mechanisms and further functional characterization of the *EfWRKY* gene family, thus providing a resource for plant breeding and genetic engineering.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13112102/s1>, Supplemental_Table S1. Protein sequences of 89 *EfWRKYs* for Phylogenetic analysis; Supplemental_Table S2. Detailed physiochemical properties of 89 *EfWRKY* proteins; Supplemental_Table S3. Secondary structure of 89 *EfWRKY* proteins; Supplemental_Table S4. Detailed information of motifs found in 89 *EfWRKYs*.; Supplemental_Table S5. Protein sequences of 83 *OsWRKY* for Phylogenetic analysis; Supplemental_Table S6. Analysis of 89 *EfWRKY*

genes cis-acting element sequences (First 2000 bp of gene sequence); Supplemental_Table S7. Cis-acting elements and functions of 89 EfWRKY genes (2000 bp before promoter); Supplemental_Table S8. Gene Duplication (KaKs values); Supplemental_Table S9. one-to-one_orthologous_relationships; Supplemental_Table S10. 89 EfWRKY genes_GO; Supplemental_Table S11. Background-GO; Supplemental_Table S12. Result_GO.

Author Contributions: H.C. and X.L. conceived and designed the experimental protocol; H.C. and D.L. performed the exploratory analysis and data compilation; H.C. and X.L. wrote the article; supervision, F.L., Y.D. and Y.F.; project administration, F.L.; funding acquisition, F.L. All authors have read and agreed to the published version of the manuscript.

Funding: Major Science and Technology Projects in Yunnan Province (202202AE090021).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the Supplementary Material.

Acknowledgments: We thank all the authors for their suggestions and critical comments on the manuscript. We also thank the reviewers for their insightful comments on this manuscript.

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

References

- Ji, Q.; Zhang, Y.; Deng, J. Research Progress of Frost on Sugarcane. *Sugar Crop. China* **2018**, *40*, 74–77.
- Gui-Hua, W. Value of Sugarcane as Feed and Its Developmental Prospect of Industrialization. *Sugar Crop. China* **2004**, 54–56.
- Gentile, A.; Dias, L.I.; Mattos, R.S.; Ferreira, T.H.; Menossi, M. MicroRNAs and drought responses in sugarcane. *Front. Plant Sci.* **2015**, *6*. [[CrossRef](#)] [[PubMed](#)]
- Zhang, Y.; Yang, S.; Dao, J.; Deng, J.; Shahzad, A.N.; Fan, X.; Li, R.; Quan, Y.; Bukhari, S.A.H.; Zeng, Z. Drought-induced alterations in photosynthetic, ultrastructural and biochemical traits of contrasting sugarcane genotypes. *PLoS ONE* **2020**, *15*, e0235845. [[CrossRef](#)]
- Shengsheng, L.; Weiwei, L.; Haiqiu, W.; Hongliang, H.; Yaqin, J.; Liqiu, T. Research Progress of Drought Stress on Sugarcane Growth. *China Trop. Agric.* **2020**, *5*, 70–73.
- Haibi, L.; Yiyun, G.; Ronghua, Z.; Jinju, W.; Rongzhong, Y.; Xiaoqiu, Z.; Rui, L.; Hui, Z.; Xihui, L. Research Progress on Drought Resistance and Drought-Resistant Breeding of Sugarcane. *Mol. Plant Breed.* **2019**, *17*, 3406–3415.
- Chao-Hua, X.; Xin, L.; Xin-Long, L.; Huo-Sheng, S.; Li, M.; Jun, M.; Xiu-Qin, L.; Xu-Juan, L.; Qing, C. Exploration and Phenotypic Traits Analysis of *Erianthus fulvus*. *J. Plant Genet. Resour.* **2014**, *15*, 962–966.
- Paul, J.R.; Imre, E.S.; Patricia, R.; Qingxi, J.S. WRKY transcription factors. *Trends Plant Sci.* **2010**, *15*, 247–258.
- Thomas, E.; Paul, J.R.; Silke, R.; Imre, E.S. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* **2000**, *5*, 199–206.
- Rushton, P.J.; Macdonald, H.; Huttly, A.K.; Lazarus, C.M.; Hooley, R. Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of α -Amy2 genes. *Plant Mol. Biol.* **1995**, *29*, 691–702. [[CrossRef](#)]
- Eulgem, T.; Rushton, P.J.; Schmelzer, E.; Hahlbrock, K.; Somssich, I.E. Early nuclear events in plant defence signalling: Rapid gene activation by WRKY transcription factors. *Embo J.* **1999**, *18*, 4689–4699. [[CrossRef](#)] [[PubMed](#)]
- Rushton, P.J.; Torres, J.T.; Parniske, M.; Wernert, P.; Hahlbrock, K.; Somssich, I.E. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *Embo J.* **1996**, *15*, 5690–5700. [[CrossRef](#)] [[PubMed](#)]
- Berg, J.M.; Shi, Y. The galvanization of biology: A growing appreciation for the roles of zinc. *Science* **1996**, *271*, 1081–1085. [[CrossRef](#)]
- Cormack, R.S.; Eulgem, T.; Rushton, P.J.; Kochner, P.; Hahlbrock, K.; Somssich, I.E. Leucine zipper-containing WRKY proteins widen the spectrum of immediate early elicitor-induced WRKY transcription factors in parsley. *Biochim. Biophys. Acta* **2002**, *1576*, 92–100. [[CrossRef](#)]
- 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 β -amylase from sweet potato. *Mol. Gen. Genet.* **1994**, *244*, 563–571. [[CrossRef](#)] [[PubMed](#)]
- Chen, C.; Chen, X.; Han, J.; Lu, W.; Ren, Z. Genome-wide analysis of the WRKY gene family in the cucumber genome and transcriptome-wide identification of WRKY transcription factors that respond to biotic and abiotic stresses. *BMC Plant Biol.* **2020**, *20*, 443. [[CrossRef](#)] [[PubMed](#)]
- Suhas, G.K.; Ranjit, S.G.; Ashutosh, R.; Moin, A.; Major, S.; Bijendra, S. Genome wide expression analysis of WRKY genes in tomato (*Solanum lycopersicum*) under drought stress. *Plant Gene* **2018**, *13*, 8–17.
- Waqas, M.; Azhar, M.T.; Rana, I.A.; Azeem, F.; Ali, M.A.; Nawaz, M.A.; Chung, G.; Atif, R.M. Genome-wide identification and expression analyses of WRKY transcription factor family members from chickpea (*Cicer arietinum* L.) reveal their role in abiotic stress-responses. *Genes Genom.* **2019**, *41*, 467–481. [[CrossRef](#)]

19. Skibbe, M.; Qu, N.; Galis, I.; Baldwin, I.T. Induced plant defenses in the natural environment: *Nicotiana attenuata* WRKY3 and WRKY6 coordinate responses to herbivory. *Plant Cell* **2008**, *20*, 1984–2000. [[CrossRef](#)]
20. Grunewald, W.; Karimi, M.; Wiczorek, K.; Van de Cappelle, E.; Wischnitzki, E.; Grundler, F.; Inze, D.; Beeckman, T.; Gheysen, G. A role for AtWRKY23 in feeding site establishment of plant-parasitic nematodes. *Plant Physiol.* **2008**, *148*, 358–368. [[CrossRef](#)]
21. Zou, X.; Seemann, J.R.; Neuman, D.; Shen, Q.J. A WRKY gene from creosote bush encodes an activator of the abscisic acid signaling pathway. *J. Biol. Chem.* **2004**, *279*, 55770–55779. [[CrossRef](#)] [[PubMed](#)]
22. Yuping, Q.; Diqiu, Y. Over-expression of the stress-induced *OsWRKY45* enhances disease resistance and drought tolerance in *Arabidopsis*. *Environ. Exp. Bot.* **2008**, *65*, 35–47.
23. Jiang, Y.; Deyholos, M.K. Functional characterization of *Arabidopsis* NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. *Plant Mol. Biol.* **2009**, *69*, 91–105. [[CrossRef](#)] [[PubMed](#)]
24. Lagacé, M.; Matton, D.P. Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*. *Planta* **2004**, *219*, 185–189. [[CrossRef](#)] [[PubMed](#)]
25. Krassimira, S.A.; Conger, B.V. Isolation of two somatic embryogenesis-related genes from orchardgrass (*Dactylis glomerata*). *Plant Sci.* **2002**, *162*, 301–307.
26. Robatzek, S.; Somssich, I.E. Targets of AtWRKY6 regulation during plant senescence and pathogen defense. *Gene. Dev.* **2002**, *16*, 1139–1149. [[CrossRef](#)]
27. Robatzek, S.; Somssich, I.E. A new member of the *Arabidopsis* WRKY transcription factor family, AtWRKY6, is associated with both senescence and defence-related processes: AtWRKY6 is associated with senescence and defence. *Plant J.* **2001**, *28*, 123–133. [[CrossRef](#)]
28. Qian, Z.; Li, X.; He, L.; Gu, S.; Shen, Q.; Rao, X.; Zhang, R.; Di, Y.; Xie, L.; Wang, X.; et al. EfGD: The *Erianthus fulvus* genome database. *Database* **2022**, *2022*, baac076. [[CrossRef](#)]
29. Mistry, J.; Finn, R.D.; Eddy, S.R.; Bateman, A.; Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **2013**, *41*, e121. [[CrossRef](#)]
30. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
31. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, w202–w208. [[CrossRef](#)]
32. Chengjie, C.; Hao, C.; Yi, Z.; Hannah, R.T.; Margaret, H.F.; Yehua, H.; Rui, X. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202.
33. He, C.; Teixeira, D.S.J.A.; Tan, J.; Zhang, J.; Pan, X.; Li, M.; Luo, J.; Duan, J. A Genome-Wide Identification of the WRKY Family Genes and a Survey of Potential WRKY Target Genes in *Dendrobium officinale*. *Sci. Rep.* **2017**, *7*, 9200. [[CrossRef](#)]
34. Wang, H.; Hao, J.; Chen, X.; Hao, Z.; Wang, X.; Lou, Y.; Peng, Y.; Guo, Z. Overexpression of rice WRKY89 enhances ultraviolet B tolerance and disease resistance in rice plants. *Plant Mol. Biol.* **2007**, *65*, 799–815. [[CrossRef](#)] [[PubMed](#)]
35. Chi-Yeol, K.; Kieu, T.X.V.; Cong, D.N.; Dong-Hoon, J.; Sang-Kyu, L.; Manu, K.; Sung-Ryong, K.; Su-Hyun, P.; Ju-Kon, K.; Jong-Seong, J. Functional analysis of a cold-responsive rice WRKY gene, *OsWRKY71*. *Plant Biotechnol. Rep.* **2016**, *10*, 13–23.
36. Taehoon, K.; Kiyoon, K.; Suk-Hwan, K.; Gynheung, A.; Nam-Chon, P. *OsWRKY5* Promotes Rice Leaf Senescence via Senescence-Associated NAC and Abscisic Acid Biosynthesis Pathway. *Int. J. Mol. Sci.* **2019**, *20*, 4437.
37. Zhou, C.; Lin, Q.; Lan, J.; Zhang, T.; Liu, X.; Miao, R.; Mou, C.; Nguyen, T.; Wang, J.; Zhang, X.; et al. WRKY Transcription Factor *OsWRKY29* Represses Seed Dormancy in Rice by Weakening Abscisic Acid Response. *Front. Plant Sci.* **2020**, *11*, 691. [[CrossRef](#)]
38. Wu, X.; Shiroto, Y.; Kishitani, S.; Ito, Y.; Toriyama, K. Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing *OsWRKY11* under the control of HSP101 promoter. *Plant Cell Rep.* **2009**, *28*, 21–30. [[CrossRef](#)]
39. Xiao, J.; Cheng, H.; Li, X.; Xiao, J.; Xu, C.; Wang, S. Rice WRKY13 regulates cross talk between abiotic and biotic stress signaling pathways by selective binding to different cis-elements. *Plant Physiol.* **2013**, *163*, 1868–1882. [[CrossRef](#)]
40. Raineri, J.; Wang, S.; Peleg, Z.; Blumwald, E.; Chan, R.L. The rice transcription factor *OsWRKY47* is a positive regulator of the response to water deficit stress. *Plant Mol. Biol.* **2015**, *88*, 401–413. [[CrossRef](#)]
41. Holub, E.B. The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* **2001**, *2*, 516–527. [[CrossRef](#)] [[PubMed](#)]
42. Ding, X.; Hou, X.; Xie, K.; Xiong, L. Genome-wide identification of BURP domain-containing genes in rice reveals a gene family with diverse structures and responses to abiotic stresses. *Planta* **2009**, *230*, 149–163. [[CrossRef](#)] [[PubMed](#)]
43. Ma, K.; Xiao, J.; Li, X.; Zhang, Q.; Lian, X. Sequence and expression analysis of the C3HC4-type RING finger gene family in rice. *Gene* **2009**, *444*, 33–45. [[CrossRef](#)] [[PubMed](#)]
44. Yue, G.; Hu, X.; He, Y.; Yang, A.; Zhang, J. Identification and characterization of two members of the FtsH gene family in maize (*Zea mays* L.). *Mol. Biol. Rep.* **2010**, *37*, 855–863. [[CrossRef](#)]
45. Yang, L.; Zhu, H.; Guo, W.; Zhang, T. Molecular cloning and characterization of five genes encoding pentatricopeptide repeat proteins from Upland cotton (*Gossypium hirsutum* L.). *Mol. Biol. Rep.* **2010**, *37*, 801–808. [[CrossRef](#)]
46. Xu, P.; Zhu, H.L.; Xu, H.B.; Zhang, Z.Z.; Zhang, C.Q.; Zhang, L.X.; Ma, Z.Q. Composition and phylogenetic analysis of wheat cryptochrome gene family. *Mol. Biol. Rep.* **2010**, *37*, 825–832. [[CrossRef](#)]
47. Wang, Y.; Deng, D.; Bian, Y.; Lv, Y.; Xie, Q. Genome-wide analysis of primary auxin-responsive Aux/IAA gene family in maize (*Zea mays* L.). *Mol. Biol. Rep.* **2010**, *37*, 3991–4001. [[CrossRef](#)]

48. Gollan, P.J.; Bhawe, M. Genome-wide analysis of genes encoding FK506-binding proteins in rice. *Plant Mol. Biol.* **2010**, *72*, 1. [[CrossRef](#)]
49. Huang, J.; Zhao, X.; Yu, H.; Ouyang, Y.; Wang, L.; Zhang, Q. The ankyrin repeat gene family in rice: Genome-wide identification, classification and expression profiling. *Plant Mol. Biol.* **2009**, *71*, 207–226. [[CrossRef](#)]
50. Dai, X.; You, C.; Wang, L.; Chen, G.; Zhang, Q.; Wu, C. Molecular characterization, expression pattern, and function analysis of the OsBC1L family in rice. *Plant Mol. Biol.* **2009**, *71*, 469–481. [[CrossRef](#)]
51. Ma, Z.; Li, W.; Wang, H.; Yu, D. WRKY transcription factors WRKY12 and WRKY13 interact with SPL10 to modulate age-mediated flowering. *J. Integr. Plant Biol.* **2020**, *62*, 1659–1673. [[CrossRef](#)] [[PubMed](#)]
52. Zhong-Qi, F.; Xiao-Li, T.; Wei, S.; Jian-Fei, K.; Wang-Jin, L.; Jian-Ye, C.; Hikmet, B. BrWRKY65, a WRKY Transcription Factor, Is Involved in Regulating Three Leaf Senescence-Associated Genes in Chinese Flowering Cabbage. *Int. J. Mol. Sci.* **2017**, *18*, 1228.
53. Lei, R.; Li, X.; Ma, Z.; Lv, Y.; Hu, Y.; Yu, D. Arabidopsis WRKY2 and WRKY34 transcription factors interact with VQ20 protein to modulate pollen development and function. *Plant J. Cell Mol. Biol.* **2017**, *91*, 962–976. [[CrossRef](#)] [[PubMed](#)]
54. Wang, P.; Yue, C.; Chen, D.; Zheng, Y.; Zhang, Q.; Yang, J.; Ye, N. Genome-wide identification of WRKY family genes and their response to abiotic stresses in tea plant (*Camellia sinensis*). *Genes Genom.* **2019**, *41*, 17–33. [[CrossRef](#)] [[PubMed](#)]
55. Lee, H.; Cha, J.; Choi, C.; Choi, N.; Ji, H.; Park, S.R.; Lee, S.; Hwang, D. Rice WRKY11 Plays a Role in Pathogen Defense and Drought Tolerance. *Rice* **2018**, *11*, 5. [[CrossRef](#)] [[PubMed](#)]
56. Yokotani, N.; Sato, Y.; Tanabe, S.; Chujo, T.; Shimizu, T.; Okada, K.; Yamane, H.; Shimono, M.; Sugano, S.; Takatsuji, H.; et al. WRKY76 is a rice transcriptional repressor playing opposite roles in blast disease resistance and cold stress tolerance. *J. Exp. Bot.* **2013**, *64*, 5085–5097. [[CrossRef](#)]
57. Tripathi, P.; Rabara, R.C.; Langum, T.J.; Boken, A.K.; Rushton, D.L.; Boomsma, D.D.; Rinerson, C.I.; Rabara, J.; Reese, R.N.; Chen, X.; et al. The WRKY transcription factor family in *Brachypodium distachyon*. *BMC Genom.* **2012**, *13*, 270. [[CrossRef](#)]
58. van Verk, M.C.; Pappaioannou, D.; Neeleman, L.; Bol, J.F.; Linthorst, H.J.M. A Novel WRKY transcription factor is required for induction of PR-1a gene expression by salicylic acid and bacterial elicitors. *Plant Physiol.* **2008**, *146*, 1983–1995. [[CrossRef](#)]
59. Tao, X.; Chengjie, C.; Chuhao, L.; Jiarou, L.; Chaoyang, L.; Yehua, H. Genome-wide investigation of WRKY gene family in pineapple: Evolution and expression profiles during development and stress. *BMC Genom.* **2018**, *19*, 490.