

Exploring the Role of WRKY Transcription Factor in the Growth and Development of *Bletilla striata* based on Bioinformatics Analysis

Qingqing Li (✉ 2744627883@qq.com)

Zunyi Medical College: Zunyi Medical University

Ceyin Huang

Zunyi Medical College: Zunyi Medical University

Chaobo Liu

Zunyi Medical College: Zunyi Medical University

Xueyan Jia

Zunyi Medical College: Zunyi Medical University

Weie Wen

Zunyi Medical College: Zunyi Medical University

Lin Li

Zunyi Medical College: Zunyi Medical University

Yihuai He

Affiliated Hospital of Zunyi Medical University

Delin Xu



Zunyi Medical College: Zunyi Medical University

Research Article

Keywords: *Bletilla striata*, bioinformatics, functional exploring, SSR marker mining, WRKY gene family

Posted Date: November 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1023129/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

WRKY type transcription factors (TFs) play crucial roles in the growth and development of plants. However, a comprehensive analysis of the WRKY family members in a valuable Chinese herbal orchid, *Bletilla striata*, or in other orchids, is limited. In this study, WRKY gene family was screened out from the transcriptome data of *Bletilla striata* by bioinformatics method. The 29 WRKY TFs that were identified from the *B. striata* genome and named *BsWRKY1* to *BsWRKY29* were divided into three clades: group I (involving 8 WRKY sequences), group II (18) and group III (3), in which group I was further divided into 5 subgroups: I-a (involving 1 WRKY sequences), I-b (5), and I-c (3), I-d (7), I-e (3). EST-SSR marker mining test showed that 10 markers could be stably amplified with obvious polymorphisms among 4 landraces. Our data suggest that *BsWRKY* genes may work together to regulate plant growth and development. In different subcellular locations, *BsWRKY* genes not only played its own functions, but also coordinated the regulation of the whole life activities. Taken together, these results provided a theoretical basis for further studies on the gene functions and regulatory mechanisms of what in *B. striata*.

Introduction

Transcription factors (TFs) are proteins that can bind to the specific sequence of gene, making gene express at a specific intensity at a specific period and tissue. WRKY transcription factors are important regulatory factors for plant growth and development in higher plants. They are involved in biological stress responses and abiotic stress responses such as drought, salt injury, pests and diseases, nutrient deficiency, high temperature and cold injury. Chen et al found that WRKY regulates fruit growth and development in jujube (Chen et al., 2019). WRKY combines with *NPR1* to regulate gene expression and then activate plant defense response (Wang et al., 2006). Cheng et al demonstrated that *WRKY13*, *WRKY45-2* and *WRKY42* form regulatory cascades that are involved in rice blast resistance (Cheng et al., 2015). Xu et al used gene microarray technology to find that *SbWRKY14*, *SbWRKY32* and *SbWRKY39* were all involved in drought stress response of sorghum (Xu, 2021). *ClWRKY47* of lemon and *CsWRKY47* of sweet orange can be induced by high salinity, drought and low temperature stress (Shen, 2021). The *ZmWRKY102* transcription factor in maize can improve the drought resistance of plants (Li, 2015). The WRKY transcription factor of *Cinnamomum kanehirae* is involved in response to plant drought and low temperature stress (Zhao, 2020). WRKY transcription factor family is also involved in plant nutrient stress and plant hormone signal transduction (Jing, 2021; Bu, 2020). The conserved WRKYGQK domain, which consists of about 60 amino acids in the DNA binding domain, is the most conserved structural feature of the WRKY transcription factor family. The downstream of this family has a cis-acting element W-box [(T/C) TGAC (C/T)] structure, which regulates the transcription of target genes and responds to adversity stress by combining with specific W-box (Wang, 2020). The C-terminal is an absolutely conserved domain composed of WRKYGQK, while the N-terminal contains a conserved domain composed of zinc finger protein. It is classified three types according to the number and structure of domains. The group I typically contains two WRKY domains including a C₂H₂ (CX₄-5CX₂₂-23HXH) motif, while group II and group III are characterized by a single WRKY domain. Group I also contains a C₂H₂ zinc-finger motif which can be further divided into five subgroups (I-a-I-e) based on the phylogeny of the WRKY domains, whereas group II contains a zinc-finger motif ending with C₂HC (CX₇CX₂₃HCX), which is only found in higher plants (Chen et al., 2020).

Bletilla striata (Thunb.) Reich, f. is a perennial herb of the orchid family, which contains chemicals of phenolic acids, dihydrophenanthrene, bibenzyl and other components with medications of anti-tumor, hemostasis, antibacterial, anti-inflammatory, promoting wound healing and plasma substitutes (Tang, 2014). It is also used in food, cosmetics and other industries (Qian et al., 2015; Zhou et al., 2020). Abundant secondary metabolites are the material basis of the pharmacological action of the *B. striata*. Therefore, from the perspective of gene family, it is of great significance to analyze the synthesis pathway of secondary metabolites of *B. striata* and explore functional genes, which helps to

understand the regulatory mechanism of effective components synthesis, and finally enhance the producing of secondary metabolites by plants. WRKY TFs have been studied in many plants such as peach (Yanbing et al., 2020), apple (Gu, 2015), pepper (Diao, 2015), melon (Ma, 2017), castor bean (Zou, 2013), *Hordeum vulgare* (Jiang et al., 2021), *Prunella vulgaris* (Zhu et al., 2020), etc., but little is known about *WRKY* gene in *B. striata*. Based on transcriptome data of *B. striata*, screening and identification WRKY TFs, and then analyze their genetic information, conservative domain, evolutionary relationships and functions, etc.. In addition, SSR molecular markers were mined to classify the function of WRKY TFs in *B. striata*, so as to provide a reference for further exploration of its function in the regulation of secondary metabolite synthesis.

1. Materials And Methods

1.1 Materials

The *B. striata* capsules were collected from the *B. striata* Germplasm Garden of Zunyi Medical University, Xinpu District, Zunyi City of Guizhou Province, China (27°42'N, 107°01'E), and the seeds were induced for suspension culture for a total of 45 days (Pan et al., 2020). The samples were randomly sampled every 3 days since from the callus was induced (3 replicates were taken at each time point), and the total RNA of each sample was extracted by liquid nitrogen grind. The RNA of each sample was mixed as one with equal amount to perform the subsequent transcriptome sequencing by using Iso-seq of PacBio platform (Li et al., 2020). The *WRKY* gene sequence of *B. striata* was screened out from the sequencing results for the following analysis.

1.2 Methods

1.2.1 *WRKY* gene family identification

The total RNA of the mixed samples was retro-transcribed into cDNA for RNA-seq sequencing. The sequencing was performed on platform of PacBio with the parameters of Iso-Seq. And the resulting data were de novo assembled by using Trinity software to finally obtain the transcriptome data. The online software Pfam and NCBI blast were employed to do the annotation. The sequences annotated as *WRKY* genes were screened to identify the conserved domains. All obtained protein sequences were examined for the presence of *WRKY* (PF03106) domains by using the Hidden Markov Model of Pfam, SMART and InterPro tools. After eliminating incomplete sequences, candidate sequences were obtained for the following analysis.

1.2.2 Physical and chemical properties exploring

The ORF Finder of NCBI was used to find out the open reading frame (ORF) of all *BsWRKY* candidate. Online server ExPASy was used to characterize the physical and chemical properties of WRKY transcription factors, such as the protein molecule weight, amino acid size, isoelectric point, instability index, GRVAY index and so on. Protein secondary structure was predicted using the online server SOPMA.

1.2.3 Analysis of subcellular localization and conserved domains

The conserved domains of the detected *WRKY* members were analyzed by using the CD-search function of NCBI. A conservative motif analysis was conducted through the MEME, in which the number of recognized motifs was set to

10. The WoLFPSORT was used to predict the subcellular localization of the WRKY proteins.

1.2.4 Signal peptide, transmembrane structure prediction and promoter analysis

Signal P 4.0 Server was used to analyze the signal peptide and TMHMM-2.0 was used to predict the transmembrane domain. The cis-acting elements of the upstream sequence (2000 bp) of the *WRKY* genes' promoter were analyzed by using the online Plant CARE, and the results were visualized by using TBTools.

1.2.5 Evolutionary analysis

The *Arabidopsis thaliana* *WRKY* sequences were obtained from the transcription factor database of Plant TFDB, *Dendrobium catenatum* *WRKY* sequences were obtained from NCBI and the MEGA-X software was applied to construct a phylogenetic tree with the *WRKY* gene family members *via* the Neigh-Joining method (NJ), and the Bootstrap value was set to 1000.

1.2.6 Functional analysis of the *BsWRKYs* genes

The Gene Ontology (GO) was performed to do GO functional classification. The mission of the GO Consortium is to develop a comprehensive, computational model of biological systems, ranging from the molecular to the organism level, across the multiplicity of species in the tree of life. The biological pathways of the *WRKY* of *B. striata* were mapped to the reference pathways in KEGG. Based on the molecular functions and biological pathways of diseases in the KEGG database, the analysis results were used to mine biologically significant information. KEGG is an encyclopedia of genes and genomes used to assign functional meaning to gene/protein elements at the molecular and higher levels. The differentially expressed cytoplasmic and nuclear proteins were matched with the KEGG pathway database to generate the predicted pathway (Ramdas et al., 2019).

1.2.7 SSR detection and verification

The NWISRL was used to detect the candidate SSR sites of *BsWRKYs*. Then, primers of each site were designed by using DNAMAN 6.0 program. The SSR candidates were verified on four landraces of *B. striata*, which were collected from Zheng'an, Chongqing, Xiuwen and Anhui, by coupling conducts of PCR and PAGE. PCR reaction system was volume 10 μ L in total, containing 1.5 μ L DNA template with concentration of 50 ng/ μ L, 6 μ L 2 \times PCR MIX, 0.75 μ L primer each with concentration of 10 μ mol/L, and 1 μ L ddH₂O. Amplification conditions were set as: pre-denaturation at 95 $^{\circ}$ C for 5 min; denaturation at 95 $^{\circ}$ C for 30 s, annealing at 52 $^{\circ}$ C for 30 s, extension at 72 $^{\circ}$ C for 60 s, 30 cycles; extension at 72 $^{\circ}$ C for 5 min. A 10% polyacrylamide gel was used to separate the amplified products. The electrophoresis instrument was a PowerPac stabilized steady flow electrophoresis instrument. The electrophoresis buffer was 1 \times TBE, the constant voltage was 150 V, and the time was 150 minutes. After silver nitrate staining, the bands were observed and photographed.

2. Results

2.1 Physical and chemical properties of WRKY protein

A total of 135 sequences annotated as *WRKY* gene were preliminarily obtained from the transcriptome data of *B. striata*. After deletions of the sequences without typical WRKY domain and incomplete sequences, 29 *WRKY* sequences were finally reserved and renamed as *BsWRKY 1-29*. Physical and chemical properties analysis showed that the protein size of the *BsWRKY* members was between 159-703 aa, and the molecular weight was between 17546.1-76820.1 Da (Table 1). The theoretical isoelectric point of proteins ranged from 4.48 to 9.94, 11 of them were basic proteins with isoelectric point greater than 7.5, 13 of them were acidic proteins with isoelectric point less than 6.5, and 5 of them were neutral between 6.5 and 7.5. These results indicated that most proteins of *BsWRKYs* were acidic. The instability coefficients of the 29 *WRKY* proteins were all greater than 40, while fat index were less than 100, and GRVAY values were negative, indicating that the *WRKY* transcription factor family of *B. striata* was an unstable hydrophilic protein. The predicted secondary structure of the protein showed that α -helix, β -folding and elongation accounted for 21.07%, 4.07% and 11.84%, respectively, and random coil accounted for 63.02% (Table 1). Among the 29 proteins, the β -turn was normally more than the α -helix except in *BsWRKY20*.

Table 1

Physical and chemical properties of *WRKY* protein of *Bletilla striata*

Gene name	AA	MW/Da	PI	Instability index	GRAVY	Alpha helix	Extended strand	Beta turn
<i>BsWRKY1</i>	307	34243	5.4	43.37	-0.651	24.76%	10.10%	3.58%
<i>BsWRKY2</i>	481	52166.7	7.01	46.65	-0.672	10.19%	14.35%	3.74%
<i>BsWRKY3</i>	406	44754.8	5.5	47.5	-0.695	25.62%	6.40%	4.43%
<i>BsWRKY4</i>	702	76105.4	5.67	48	-0.627	11.97%	11.11%	3.42%
<i>BsWRKY5</i>	500	54510.7	8.97	48.21	-0.752	12.00%	12.80%	3.60%
<i>BsWRKY6</i>	469	50789.1	6.98	50.9	-0.729	10.45%	14.07%	3.41%
<i>BsWRKY7</i>	185	21037.2	6.26	51.01	-0.97	24.86%	18.92%	10.27%
<i>BsWRKY8</i>	194	51538.4	9.59	52.71	-0.865	11.34%	13.92%	6.19%
<i>BsWRKY9</i>	335	35817.5	9.69	53.06	-0.484	21.49%	9.85%	6.87%
<i>BsWRKY10</i>	431	47054.3	6.57	54.84	-0.703	14.39%	12.30%	2.55%
<i>BsWRKY11</i>	341	36780.7	9.67	55.85	-0.484	18.18%	10.26%	6.45%
<i>BsWRKY12</i>	329	37045.9	8.57	56.08	-0.632	32.22%	10.33%	1.52%
<i>BsWRKY13</i>	347	38224.3	9.94	56.17	-0.688	25.07%	9.22%	5.48%
<i>BsWRKY14</i>	270	29189.3	5.64	56.66	-0.0743	17.04%	11.11%	3.70%
<i>BsWRKY15</i>	217	24804.6	5.91	57.12	-0.86	11.06%	15.21%	3.69%
<i>BsWRKY16</i>	233	25270	4.93	57.21	-0.642	27.90%	10.30%	3.43%
<i>BsWRKY17</i>	412	45995.3	8.62	57.49	-0.617	31.31%	8.98%	1.94%
<i>BsWRKY18</i>	595	64961.1	6.55	57.8	-0.65	28.57%	10.92%	1.51%
<i>BsWRKY19</i>	402	43491.8	8.58	59.9	-0.829	9.95%	11.44%	3.23%
<i>BsWRKY20</i>	159	17546.1	4.48	60.03	-0.82	40.25%	11.32%	8.81%
<i>BsWRKY21</i>	427	47131.5	5.25	60.59	-0.63	23.89%	6.56%	3.51%
<i>BsWRKY22</i>	174	19581.5	4.55	60.82	-0.506	21.84%	18.39%	5.17%
<i>BsWRKY23</i>	595	64961.1	6.55	62.18	-0.65	28.57%	10.92%	1.51%
<i>BsWRKY24</i>	335	35367.1	7.61	62.24	-0.354	25.37%	13.73%	3.88%
<i>BsWRKY25</i>	530	56563.4	6.36	62.43	-0.622	20.94%	10.94%	2.45%
<i>BsWRKY26</i>	703	76820.1	5.61	62.54	-0.648	13.23%	10.53%	3.27%
<i>BsWRKY27</i>	363	41768.4	8.52	63.93	-0.781	31.68%	13.50%	2.75%
<i>BsWRKY28</i>	526	57802.3	6.2	66.1	-0.657	13.50%	12.93%	3.42%
<i>BsWRKY29</i>	343	38673.7	9.76	66.63	-1.146	23.32%	12.83%	4.37%

2.2 Subcellular localization, signal peptide and transmembrane structure

The subcellular localization prediction showed that 24 WRKY proteins were all located in nucleus, except that BSWRKY16 was located in chloroplast, BSWRKY14 and BSWRKY29 were located in mitochondria, BSWRKY17 was located in vacuole, and BSWRKY28 was located in endoplasmic reticulum. Subcellular location determines its specific biological effects. WRKYs can form a net that contributes to various cytoplasmic and nuclear processes including signaling events from organelles or the cytoplasm to the nucleus (Bakshi and Oelmüller, 2014). Studies had shown that WRKY TFs on the ABA-ABA complex in the downstream chloroplast envelope, regulates seed germination and other processes, and is one of the key nodes of abscisic acid signaling pathways (Rushton et al., 2012). These results indicated that *WRKY* genes might be involved in the regulation of plant growth and development and is an important node in metabolic regulation.

2.3 Promoter cis- regulatory elements of *BsWRKY* genes

The upstream of these *BsWRKY* genes were detected for finding cis-regulatory elements, like promoter and other cis-acting elements related to hormone regulation and stress-response (Fig. 1). The result showed that the cis-regulatory elements of the promoters of *BsWRKY* genes were related to growth and development (meristem expression, specific to the endosperm, seed-specific regulation and regulates circadian rhythm), plant hormones (auxin, abscisic acid, methyl jasmonate (MeJA), gibberellin, and salicylic acid), and stress (drought, low temperature, oxygen specificity induced response element and anaerobic induced indispensable cis function adjustment). It also showed that all the 29 *BsWRKY* genes had light response elements (LRE), and 14 of them had the drought-inducibility response elements. However, the elements of cell cycle regulatory elements, elements involved in defense and stress response, flavonoid synthesis and seed germination were only existed in *BsWRKY28*, *BsWRKY7*, *BsWRKY16* and *BsWRKY24*, respectively. This not only indicated that *BsWRKY* genes are associated with plant growth, but also playing a vital role in drought stress regulatory networks. Collectively, these results indicated that *WRKY* family members participate in embryonic development, meristem growth and environmental stress regulation during the growth and development of *B. striata*.

2.4 Conservative motif of WRKY protein

A total of 10 conserved motifs were obtained by using online MEME for motif analysis of WRKY transcription factors in *B. striata* (Fig. 2A, B, C). Among them, motif3 were contained in 28 members except BSWRKY7. Motif1 and motif2 existed in 26 members, motif8 was found in 6 members, motif4 was detected from 4 members. The CD-search analysis found that motif1, motif2, and motif4 belong to WRKY domains, motif8 was a zinc finger domain relates to WRKY. Interestingly, the motif3 had no function record in the database currently, which needs to be further studied.

2.5 Conserved domains identification and evolutionary analysis

Through the online server CD-search, the structure domains of the WRKY genes family were analyzed for comparison. The results showed that 24 members of the 29 *BsWRKY* had typical WRKYGQK heptapeptide domain and W-box, but there was different degree of variation which mainly occurred in the N-terminal. Five transcription factors, i.e. *BsWRKY 7, 16, 17, 22* and *27*, had incomplete domains, like missing N or C terminus. It was speculated that the deletion may occur in the evolutionary process.

To place the evolution role and further identify the functions of *BsWRKY* genes, these 29 *WRKY* sequences from *B. striata*, 22 *WRKY* from *Arabidopsis thaliana* and 20 *WRKY* from *Dendrobium catenatum* were used to perform a phylogenetic analysis (Fig.3). The 29 *BsWRKY* transcription factors were grouped into three major groups. Group I included 6 members as *BsWRKY 4, 17, 19, 10, 25, 26 and 28*, which contains two *WRKY* conserved domains. Group III had members of *BsWRKY1, 15 and 22*, which containing only one *WRKY* domain with the zinc finger structure of type C₂HC. In addition, Group II had only one domain with the zinc finger structure of C₂H₂, which could be further divided into 5 subgroups. In which, subgroup II-a had only member of *BsWRKY12*, subgroup II-b had four as *BsWRKY 18, 23, 24 and 27*, subgroup II-c had *BsWRKY2, 5, 6, 7, 10 and 29*, subgroup II-d was consisted by six members i.e. *BsWRKY8, 9, 11, 13, 14 and 16*, subgroup II-e hold three members as *BsWRKY3, 20 and 21*.

According to the evolutionary tree, event of domain's gain or loss might occur in the process of evolution, and group I was closer the real ancestor. The evolutionary relationship of *WRKY* genes among *B. striata*, *A. thaliana* and *D. catenatum* indicated that the three may have similar roles in certain biological functions. The results showed that *BsWRKY5*, *BsWRKY12* and *BsWRKY3* had the highest similarity with *AtWRKY3*, *AtWRKY40* and *AtWRKY14*, respectively.

2.6 GO enrichment and KEGG functional cluster of *WRKY* genes

The GO and KEGG analysis results (Table 3) showed that most of *BsWRKY* genes were divided into three categories with functions of cellular components, biological processes and molecular functions, except *BsWRKY8, 14, 16, 20 and 22*. Among the biological processes, 19 genes were annotated for transcriptional regulation, *BsWRKY18* and *BsWRKY27* for cell cycle, *BsWRKY25* for DNA repair and *BsWRKY17* for meiosis prophase. In the molecular function category, 16 proteins were annotated for sequence specific DNA binding and 20 proteins for sequence specific DNA binding transcription factor activity. Among the large group of cellular components, only *BsWRKY12* was annotated into the troponin complex, and *BsWRKY25* had DNA ligase (NAD⁺) activity. KEGG analysis showed that *BsWRKY19, 20, 26 and BsWRKY28* were involved in environmental information processing, signal transduction, MAPK signal transduction pathway, environmental adaptation, plant pathogen interaction and biological system.

Table 3

GO function classification of *WRKY* protein of *B. striata*

Gene name	Biological processes				Molecular function		Cellular components	
	Transcriptional regulation	DNA repair	Cell cycle	Meiosis prophase	Sequence-specific DNA binding	Sequence-specific DNA binding transcription factor activity	DNA ligase (NAD+) activity	Troponin complex
<i>BsWRKY 1</i>	√				√	√		
<i>BsWRKY 2</i>	√				√	√		
<i>BsWRKY 3</i>	√				√	√		
<i>BsWRKY 4</i>	√				√	√		
<i>BsWRKY 5</i>	√				√	√		
<i>BsWRKY 6</i>							√	
<i>BsWRKY 7</i>	√				√	√		
<i>BsWRKY 8</i>								
<i>BsWRKY 9</i>	√				√	√		
<i>BsWRKY 10</i>	√				√	√		
<i>BsWRKY 11</i>							√	
<i>BsWRKY 12</i>	√				√	√		√
<i>BsWRKY 13</i>	√				√	√		
<i>BsWRKY 14</i>								
<i>BsWRKY 15</i>	√				√	√		
<i>BsWRKY 16</i>								
<i>BsWRKY 17</i>	√			√				
<i>BsWRKY 18</i>	√						√	
<i>BsWRKY 19</i>	√				√	√		

<i>BsWRKY</i> 20				
<i>BsWRKY</i> 21	√		√	√
<i>BsWRKY</i> 22				
<i>BsWRKY</i> 23				√
<i>BsWRKY</i> 24	√		√	√
<i>BsWRKY</i> 25		√		√
<i>BsWRKY</i> 26	√		√	√
<i>BsWRKY</i> 27	√		√	
<i>BsWRKY</i> 28				√
<i>BsWRKY</i> 29	√		√	√

2.7 EST-SSR polymorphism of *WRKY* genes in *B. striata*

The EST-SSR markers have the advantages of high polymorphism and variability, high reproducibility, accurate and rapid detection (Li et al., 2015). A total of 10 among the 29 sequences were detected with SSR sites by NWISRL, of which 3 sequences were dinucleotide repeats and 7 sequences were trinucleotide repeats. The lowest number of replicates was 5, and the highest number was 18. The primer pairs of the 10 SSR sites were designed by DNAMAN software which could be amplified stably in all the four landraces (Table A), and the length of the amplified products ranged from 100 to 200 bp (Fig. 4). These results indicated that *WRKY* gene family was probably high-conserved in different *B. striata* germplasm. These newly found SSR primers could be used as molecular markers to identify the members of *BsWRKY* gene families in different germplasm (Zhong et al., 2021).

3. Discussion

WRKY transcription factors are important transcription factors involved in growth, development and response to stress in eukaryotes, and play an important role in plant survival. Based on bioinformatics analysis, the physicochemical properties, enzyme restriction sites, conserved motifs, cis elements, evolutionary relationships and functions of *WRKY* gene family of *B. striata* were analyzed, and EST-SSR sites were detected and verified. Based on the transcriptome data of *B. striata*, a total of 29 members of *BsWRKY* gene family were screened. Among them, 24 members had typical *WRKY*GQK conserved structure and downstream W-box [(T/C) TGAC (C/T)] homeostasis element. However, five members (*BsWRKY*7, 16, 17, 22 and 27) had different degrees of deletion at the C-terminal or N-terminal, which indicated that they might be self-regulation or cross regulation between *WRKY* genes. Among them, the length of *BsWRKY*7 conservative domain was less than half, but it had elements involved in defense and stress response. GO enrichment indicated that *BsWRKY*7 has specific sequence DNA binding transcription factor activity and participates

in cell transcription regulation, which belongs to subgroups II-c in classification. Therefore, its function and binding specificity could be further investigated.

Subcellular localization showed that 24 *BsWRKY* genes played a role in the nucleus, while *BsWRKY28* was located in the endoplasmic reticulum. It had a response element for regulating the cell cycle as well as other regulatory elements such as drought, anaerobic stress, salicylic acid response and ATBP-1 binding site. KEGG analysis showed that *BsWRKY28* was involved in environmental adaptation and also belonged to the subgroups II-c. These results indicated that it was involved in the synthesis and regulation of important substances in the process of cell growth and development and played a vital role in the stress. *BsWRKY17* was located in vacuoles and had a signal peptide and transmembrane structure. It was speculated that *BsWRKY17* may function as a secreted protein, which will be secreted into cells after synthesis and then played a role. *BsWRKY16* was located in chloroplasts with seven restriction sites, and also had cis-acting elements related to flavonoid synthesis, suggesting that *BsWRKY16* was involved in regulating the synthesis of plant secondary metabolites. Previous studies have shown that *WRKY* is also involved in the synthesis of terpenoids and played a key regulatory role (Park et al., 2021). *BsWRKY14* and *BsWRKY29* played a role in mitochondria. *BsWRKY14* had auxin response element and low temperature response element. *BsWRKY29* had MYBHv1 binding site, abscisic acid response element, hypoxia specific induction element, salicylic acid response element, ATBP-1 binding site and MeJA response element. Both of them had gibberellin response element and anaerobic induction element. So we speculated that these two elements may be involved in the regulation of cell maturation and senescence, *BsWRKY29* also belonged to subclass II-c, suggesting that subgroups II-c might play an important role in stress response (Na et al., 2018). The *BsWRKY* gene family had plant hormone or stress response elements, which were tissue-specific and have synergistic effects among members. Its function is to respond to hormonal signals and stress. We speculated that all genes contain light-response elements, which affect the cell cycle and help plants cope with various stresses. Many studies show that *WRKY* gene in banana (Zheng et al., 2021), longan (Fan et al., 2017), *Artemisia annua* (Xue and Jue, 2021), rice (Li et al., 2021), barley (Viana et al., 2021) and other plants is involved in regulating the maturity and senescence of plants, promoting the accumulation of specific products, and responding to one or more abiotic stresses. Despite these similarities, plant development is a very complicated process, and *BsWRKY* genes could be directly or indirectly involved in a certain regulatory role. Therefore, it is necessary to conduct more in-depth and detailed research on these genes.

Former studies reported that *AtWRKY3* involved in biological stress response (Aboul-Maaty and Oraby, 2019), *AtWRKY40* participated in drought stress response in *A. thaliana* (Ju et al., 2019), *AtWRKY14* played a role in plant antiviral (Che et al., 2018). Therefore, we speculated that genes *BsWRKY5*, *BsWRKY12* and *BsWRKY3* might play a regulatory role in stress resistance in *B. striata*. GO enrichment classified the *BsWRKY* gene family into three categories of biological processes, molecular functions, and cell components. In which the *WRKY* gene family was involved in important processes such as cell cycle, transcriptional regulation, and meiosis. KEGG clustering also confirmed that *BsWRKY* gene family was involved in signal transduction and environmental information processing, growth and development regulation, stress response and hormone signal transduction, which inferred that *WRKY* family was an important factor in plant growth and development and response to environmental stress. EST-SSR polymorphism detection showed that the occurrence frequency of *WRKY* gene family in *B. striata* was 34.48%, which was dominated by trinucleotide repeats, indicating that the *WRKY* family in different *B. striata* varieties was conserved. This indicated that this gene family has genetic conservative. Therefore, the SSR sites can be used to identify the genetic diversity of germplasm resources, and provide some reference for the molecular-assisted breeding and genetic diversity analysis of plants, which is of great significance for the development and utilization of gene function and the evaluation of germplasm resources.

4. Conclusion

WRKY transcription factor family was named from the highly conserved region WRKY domain, which can specifically bind to the W-box in the promoter of the target gene and regulate the expression of the target gene. The *WRKY* gene family of *B. striata* was a highly conserved gene family with hereditary conservation among different species. It contained many hormone response elements and stress response elements, and had many enzyme digestion sites, which make it involved in the growth and development process of plants and the key points for regulation of the synthesis of secondary metabolites. It's one of regulators in primary and secondary metabolism and played an important role in the stress. According to the bioinformatics analysis and function characterization of each member of the *BsWRKY* family, we screened the relative sequence for each function. In different locations, WRKY not only played its own functions, but also coordinated the regulation of the whole life activities. These results provided a theoretical basis for further study of the functions and regulatory mechanisms of *WRKY* gene family of *B. striata*. As transcription factors, *WRKY* genes had different function, while some potential WRKY TFs may be to control the growth and development in *B. striata*, for instance, tubers development and biomass accumulation of *B. striata*. However, more functions and internal regulatory mechanisms need to be studied. By studying the *WRKY* gene family, gene editing technology can be used to conduct in-depth research on the members of this gene family, and further explore the functions of this gene family and the regulatory mechanisms involved in regulating plant growth, development and stress response. There will be opportunities to improve stress resistance and secondary metabolites in plants of great significance for cherished medicinal plants.

Declarations

Acknowledgments:

This research was financially supported by the National Natural Science Foundation of China (31960074), the Science and Technology Department Foundation of Guizhou Province of China (No. [2017]5733-050, [2019]-027, [2019]5657), the Special Joint Bidding Project of Zunyi Sci & Tech Bureau and Zunyi Medical University (ZSKHHZ-2020-91) and Honghuagang Sci & Tech Project of Zunyi City (ZHKHNZT [2020]04).

Conflict of interest Statement:

The authors declare that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

Author Contributions:

DX and YH conceived, supervised and writing-reviewed the manuscript, QL, CH, CL and XJ originally wrote and writing-reviewed the draft, LL and WW cofounded and co-administrated the project. All authors read and approved the final version.

References

1. Aboul-Maaty AF, Oraby AS (2019) Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. *Bulletin of the National Research Centre* 43(1):2–10
2. Bakshi M, Oelmüller R (2014) WRKY transcription factors: Jack of many trades in plants. *Plant Signal Behav* 9(2):e27700. doi: 10.4161/psb.27700

3. Bu Huahu (2020) Research Progress on Plant WRKY Transcription Factors Family Genes. *Journal of Shanxi Agricultural Sciences* 48(7):1158–1163
4. Che YM, Sun YJ, Song-Chong LU et al (2018) AtWRKY40 functions in drought stress response in *Arabidopsis thaliana*. *Plant Physiology Journal* 054(003):456–464
5. Chen X, Chen R, et al YW et al (2019) Genome-Wide Identification of WRKY Transcription Factors in Chinese jujube (*Ziziphus jujuba* Mill.) and Their Involvement in Fruit Developing, Ripening, and Abiotic Stress. *Genes* 10(5):2–17
6. Cheng H, Liu H, Deng Y et al (2015) The *WRKY45-2 WRKY13 WRKY42* transcriptional regulatory cascade is required for rice resistance to fungal pathogen. *Plant Physiol* 167(3):1087–1099
7. Chen C et al (2020) 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 Biology* 20(1). doi:10.1186/s12870-020-02625-8
8. Diao Wei-ping (2015) Genome-wide Analysis of the WRKY Transcription Factor Family in Pepper. *Acta Horticulturae Sinica* 42(11):2183–2196
9. Zhongqi F, Yuanfang Z, Fu Yali, et al (2017) Molecular Characterization of a Transcriptional Factor, MaWRKY31, and Its Regulation on *MaACS1*, *MaACO1* and *MaSAG1* [J]. *Food Science*, 38(06), 13-9
10. Gu Yan-bing (2015) Bioinformatics and Expression Analysis of the WRKY Gene Family in Apple. *Scientia Agricultura Sinica* 48(16):3221–3238
11. Jiang Liling C, Lihua N et al (2021) Bioinformatics Analyses of WRKY Family Members in *Hordeum vulgare* L. var. *nudum* Based on RNA-seq Data. *Molecular Plant Breeding* 19(02):370–379
12. Ju Ping L, Li Yu, Lili et al (2019) Cloning and Expression Analysis of *CmWRKY4* Gene in Chrysanthemum. *Molecular Plant Breeding* 17(01):29–32
13. Jing Yizhuo (2021) Research progress of WRKY transcription factor family in response to crop elemental stress. *Plant Physiol* 57(2):265–273
14. Li C, Li D, Shao F et al (2015) Molecular cloning and expression analysis of WRKY transcription factor genes in *Salvia miltiorrhiza*. *BMC Genomics* 16(1):200
15. Li L, Liu H, Wen W et al (2020) Full Transcriptome Analysis of Callus Suspension Culture System of *Bletilla striata*. *Front Genet* 11:1–12
16. Li Qi G, Xiao-yue ZHANG, Lei et al (2021) Research on transcription factors related to artemisinin biosynthesis in *Artemisia annua*. *Chinese Traditional and Herbal Drugs* 52(06):1827–1833
17. Li X (2015) Clong and analysis of Biological Functions of Transcription Factors *ZmWRKY102* from *Zea mays*. PhD thesis. Anhui Agricultural University
18. Ma Yong B, Mulan Z et al (2017) Bioinformatics Analysis of WRKY Transcription Factor Genes Family in Melon. *Genomics and Applied Biology* 36(11):4761–4769
19. Na Y, Li J, Han Z et al (2018) Research progress of jasmonate-responsive transcription factors in regulating plant secondary metabolism. *Zhongguo Zhong Yao Za Zhi* 43(5):897–903
20. Park SH, Scheffler JA, Ray JD et al (2021) Identification of simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) that are associated with the nectariless trait of *Gossypium hirsutum* L. *Euphytica* 217(4):1–17
21. Pan Y, Li L, Xiao S et al (2020) Callus growth kinetics and accumulation of secondary metabolites of *Bletilla striata* Rchb.f. using a callus suspension culture. *PLoS ONE*, 15(2), e0220084
22. Zhang Q et al (2015) Wet spinning of *Bletilla striata* polysaccharide/silk fibroin hybrid fibers. *Materials Letters* 161. doi:10.1016/j.matlet.2015.09.031

23. Rushton DL, Tripathi P, Rabara RC et al (2012) WRKY transcription factors: key components in abscisic acid signalling. *Plant Biotechnol J* 10(1):2–11. doi: 10.1111/j.1w467-7652.2011.00634.x
24. Ramdas P, Radhakrishnan AK, Abdu Sani AA et al (2019) Advancing the Role of Gamma-Tocotrienol as Proteasomes Inhibitor: A Quantitative Proteomic Analysis of MDA-MB-231 Human Breast Cancer Cells. *Biomolecules* 10(1):19. doi:10.3390/biom10010019
25. Shen Dan (2021) Cloning and expression analysis of stress response gene *WRKY47* in citrus. *Jiangsu J of Agr Sci* 37(1):129–138
26. Tang Yifei (2014) Research Progress on chemical constituents and pharmacological effects of *Bletilla striata*. *Chinese Traditional and Herbal Drugs* 45(19):2864–2872
27. Viana VE, Carlos da Maia L, Busanello C et al (2021) When rice gets the chills: comparative transcriptome profiling at germination shows WRKY transcription factor responses. *Plant Biol (Stuttg)* 1100–112. doi: 10.1111/plb.13262
28. Wang C (2020) Genome-wide analysis of the WRKY family response to salt stress and histidine kinase in *Beta vulgaris*. PhD thesis. Harbin Institute of Technology
29. Wang D, Amornsiripanitch N, Dong X (2006) A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS Pathog*, 2(11), e123
30. Xu L (2021) Expression Analysis of Group WRKY Transcription Factors in Sorghum Under Drought Stress. *Guangdong Agricultural Sciences* 48(2):11–16
31. Xin X, Dengwei J (2021) Cloning of Longan *DIWRKY25* Gene and Analysis of Its Function in Flowering Process. *Molecular Plant Breeding*, 1–13
32. Yanbing G, Zhirui J, Fumei C et al (2017) Genome-wide identification and expression analysis of the *WRKY* gene family in peach. *Acta Agriculturae Boreali-Sinica* 38(3):254–270
33. Zou Zhi (2013) Genome - wide identification and phylogenetic analysis of WRKY transcription factor family in castor bean. *Chinese Journal of Oil Crop Sciences* 35(1):036–042
34. Zhu Yun-hao CHEN, Zhi-heng DONG, Cheng-ming (2020) Transcriptome-wide Identification and Expression Profiles of WRKY Transcription Factor in *Prunella vulgaris*. *Chinese Journal of Experimental Traditional Medical Formulae* 26(20):146–152
35. Zhong H, Zhang F, Zhou X et al (2021) Genome-Wide Identification of Sequence Variations and SSR Marker Development in the Munake Grape Cultivar. *Frontiers in Ecology and Evolution* 9:1–11
36. Zheng J, Zhang Z, Tong T et al (2021) Genome-Wide Identification of WRKY Gene Family and Expression Analysis under Abiotic Stress in Barley. *Agronomy* 11(3):521
37. Zhou Jiaming Ou, Jingchun Z, Ai et al (2020) Research and development of *Panax notoginseng* whitening cosmetics. *Ginseng research* 032(001):40–42
38. Zhao N, He M, Li L et al (2020) Identification and expression analysis of WRKY gene family under drought stress in peanut (*Arachis hypogaea* L.). *PLoS ONE*, 15(4), e0231396

Table

Table 2 is not available with this version.

Figures

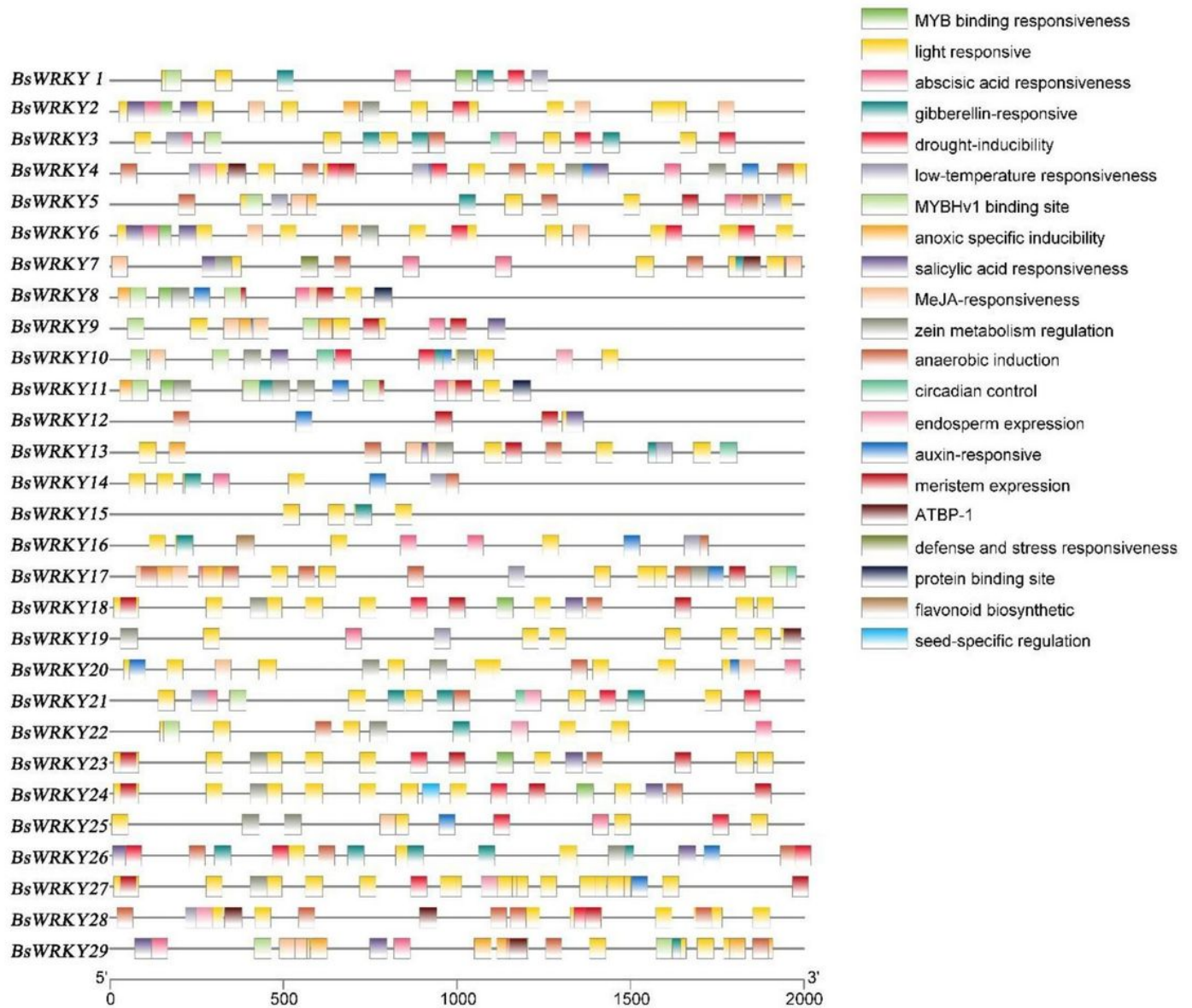


Figure 1

Cis-responsive elements in the upstream region of the initiation codon of BsWRKY genes. Different colored boxes indicated different cis-responsive elements.

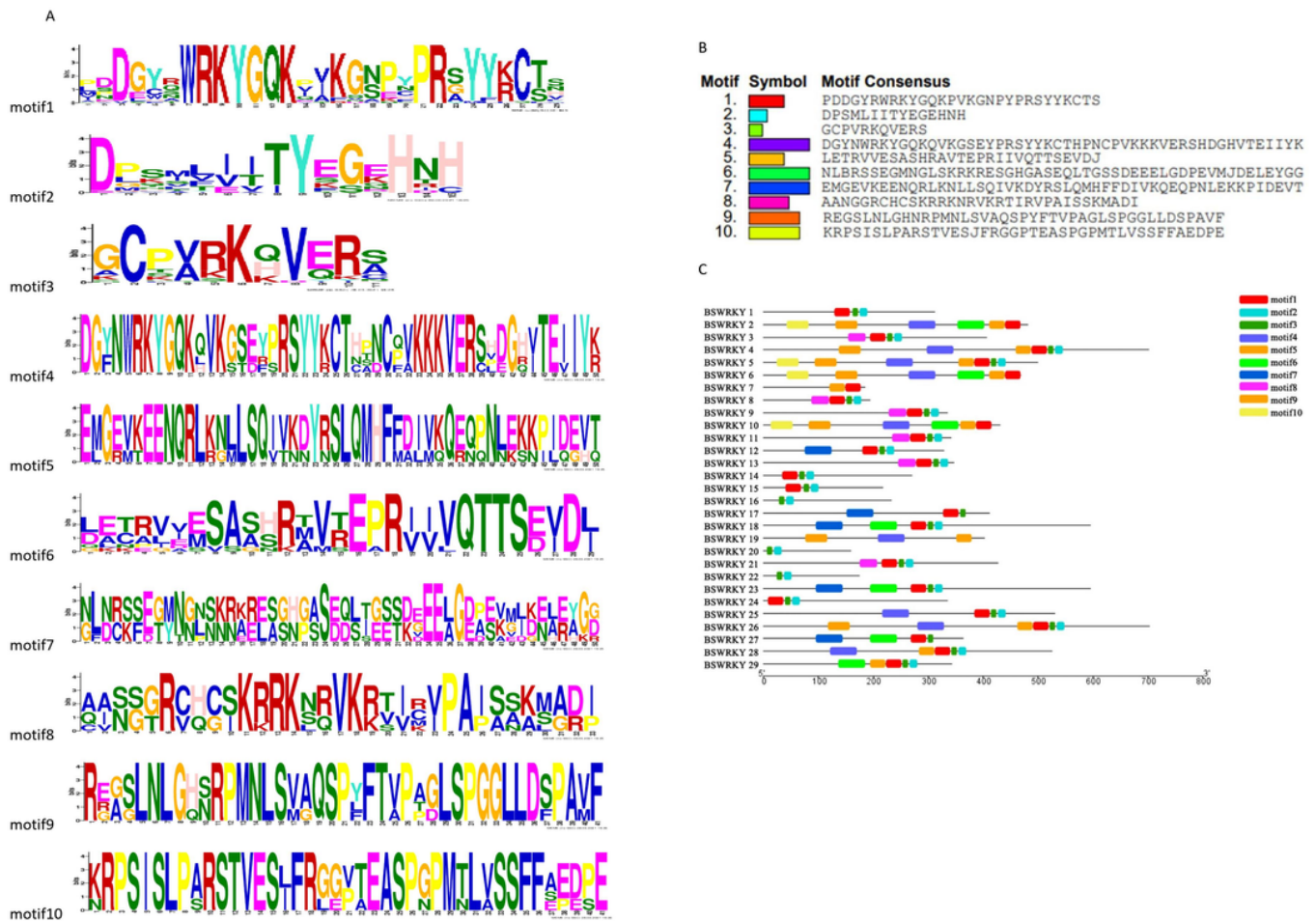


Figure 2

The size(A), sequence(B) and location(C) of the conserved motifs on WRKY protein family. The length and order of the boxes with different colors represent the actual size and location of each motif in protein sequence respectively.

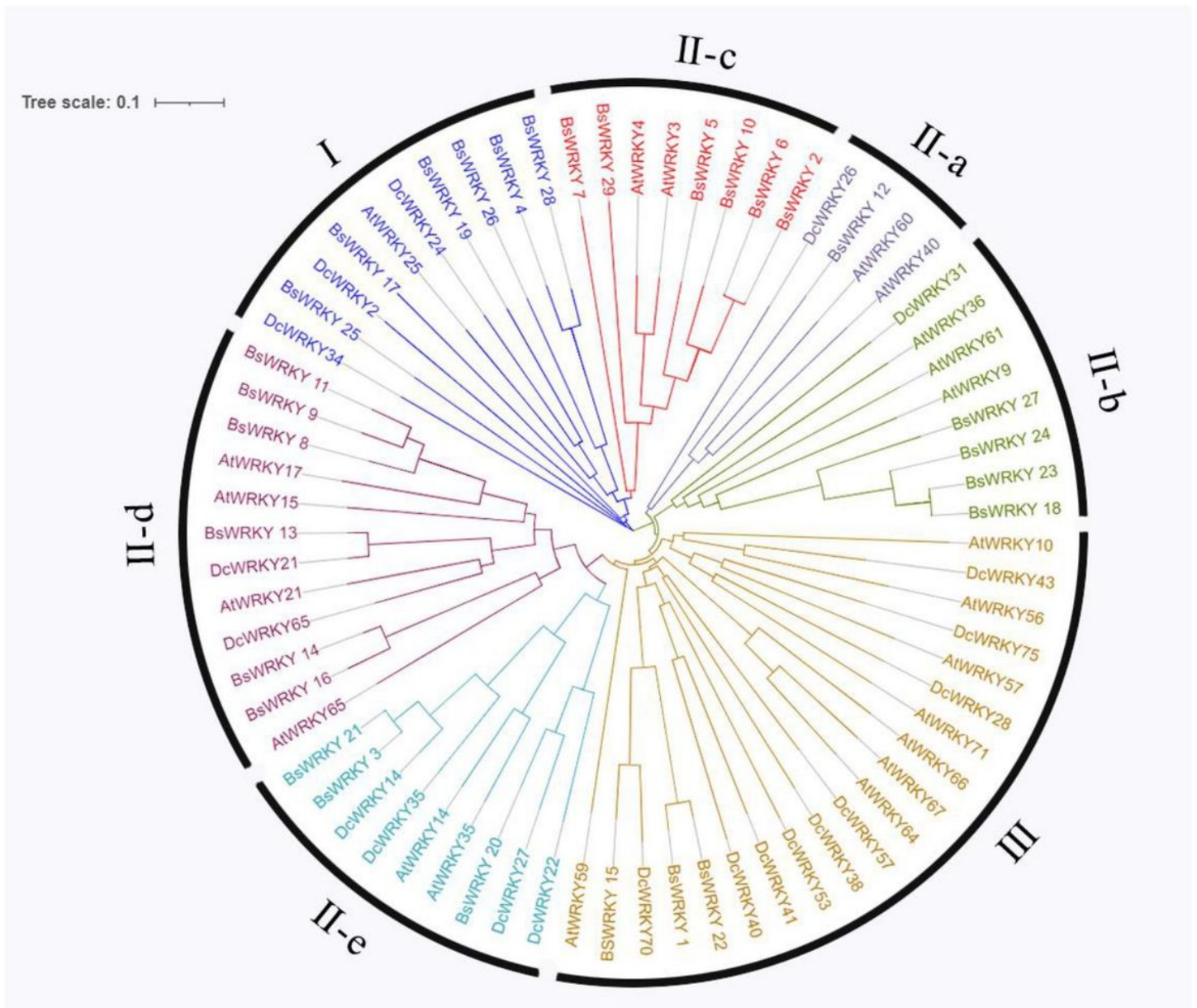


Figure 3

Phylogenetic tree of WRKY genes in *B. striata* and *A. thaliana* and *D. catenatum*.

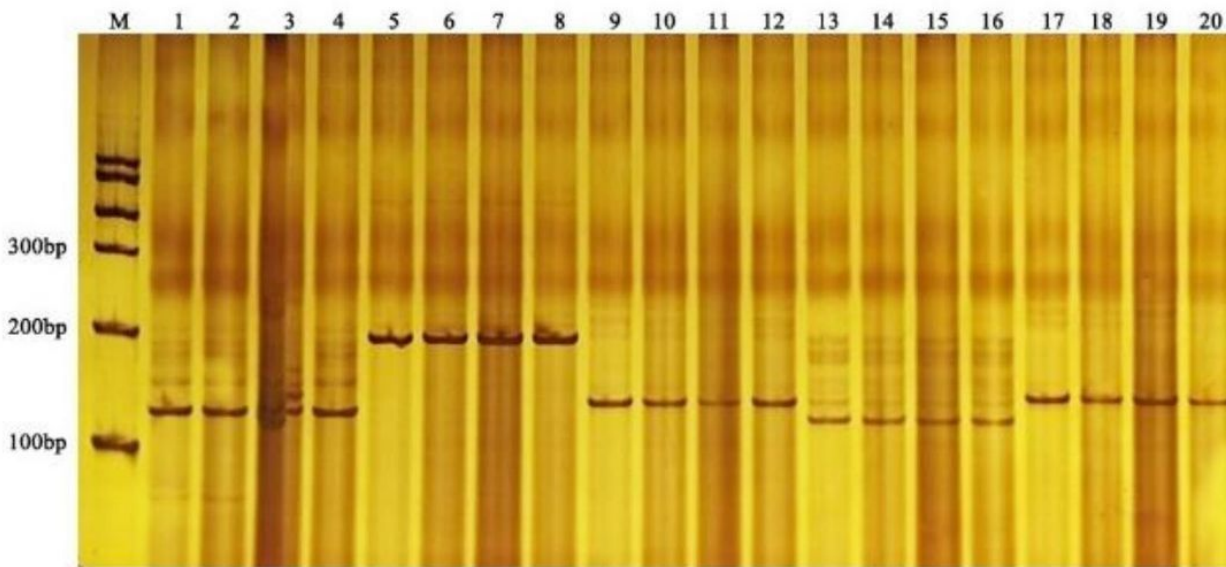
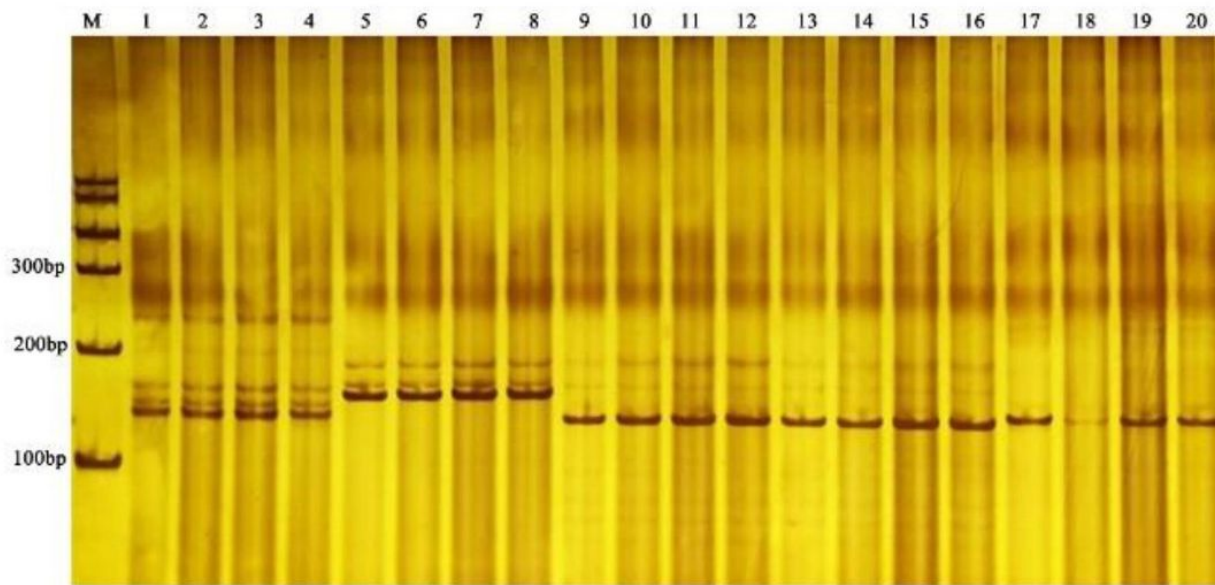


Figure 4

SSR profiles of *B. striata*.

Supplementary Files

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

- [Appendix.docx](#)