Full-Length Transcriptome-Wide Characteristic and Functional Identification of WRKY Family in *Malus sieversii* during the Valsa Canker Disease Response

Xiaojie Liu1,2, Yiheng Zhang3, Tong Zhou1,*, Xiaohuang Li1,4, Xuejing Wen1,4 and Daoyuan Zhang1,4,*

1 State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China; liuxiaojie215@mails.ucas.ac.cn (X.L.); zhoutong.cn@outlook.com (T.Z.); lixs@ms.xjb.ac.cn (X.L.); wxj-329@163.com (X.W.)
2 University of Chinese Academy of Sciences, Beijing 100049, China
3 The College of Forestry, Nanjing Forestry University, Nanjing 210037, China; Zyh17805107020@163.com
4 Turpan Eremophyte Botanical Garden, Chinese Academy of Sciences, Turpan 838008, China
* Correspondence: zhangdy@ms.xjb.ac.cn

Abstract: WRKY transcription factors are one of the largest families in plants, playing important roles in regulating plant immunity. *Malus sieversii* has abundant genetic diversity and can offer various and high-quality gene resources. In this study, 112 putative MsWRKY proteins were identified from a full-length transcriptome of *M. sieversii* during the *Valsa* canker disease (caused by *Valsa mali*). The MsWRKY proteins were phylogenetically divided into three groups (I–III). Motif compositions of the MsWRKY proteins were clustered and fifteen conserved motifs were observed. Expression pattern analysis showed that thirty-four MsWRKY transcripts strongly responded to the *V. mali* infection, demonstrating that MsWRKY transcripts might play different roles during the response. Functional identifications were subsequently conducted with transient expressions, demonstrating that MsWRKY16, MsWRKY21, MsWRKY70, MsWRKY74 and MsWRKY85 positively regulated the resistant response. Besides, the MsWRKY21, MsWRKY70 and MsWRKY85 were dramatically induced by salicylic acid (SA), methyl-jasmonate acid (MeJA) and 1-aminocyclopropane-1-carboxylate (ACC), indicating that they play important roles in the regulatory resistance of *V. mali* infection. This work provides a comprehensive understanding of the WRKY family in *M. sieversii* and will build a foundation for future research of the potential disease resistances MsWRKY transcripts.

Keywords: WRKY transcription factor; *Malus sieversii*; Valsa canker; gene expression; disease resistance

1. Introduction

WRKY transcription factor (TF) family is one of the largest families in plants and plays an important role, both positively and negatively, in transcription regulatory networks to respond to the biotic and abiotic stresses [1,2]. The first WRKY gene was identified in 1994, named *SWEET POTATO FACTOR 1* (*SPF1*) [3]. The WRKY TF family contains 74, 109 and 113 members in *Arabidopsis thaliana*, *Oryza sativa* and *Malus domestica* (*GDDH13*) [4–6].

WRKY transcription factors contain the highly conserved amino acid sequence WRKYGGQK domain in the N-terminal and zinc-finger motifs C2H2 (CX4_5CX22–23HXH) or C2HC (CX3CX23HXC) in the C-terminal, hence named WRKY TF family [7]. Recent research reported that the highly conserved heptapeptide WRKYGGQK also had other forms, which are WRKYGK, WKKYGQK, WRKYGQR and WRKYGEK [8,9]. WRKY transcription factors have sequence-specific DNA-binding activity, named W box (C/T) TGAC (T/C) element [10]. The WRKY transcription factors are classified into three main groups based on the number of WRKY domains and the type of zinc-finger [7]. WRKY proteins with two WRKY domains (in N-terminal and C-terminal, respectively) were classified into group I, while proteins with only one WRKY domain belong to group II or III. Generally,
the members of the group I and II have the C2H2 zinc-finger motif, while the members of group III contain the C2CH zinc-finger motif. Additionally, WRKY group II is divided into five subgroups, Ila–Ile [7]. Except for the WRKY domain, some WRKY transcription factors have other domains, which are B3, TIR-NBS-LRRs, kinase domain, basic nucleus location signals and leucine zippers domains [11,12].

WRKY transcription factors are central regulators in the plant innate immunity, which are microbe-associated molecular pattern-triggered immunity (MTI), pathogen-associated molecular pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) [13]. Several WRKY proteins’ phosphorylation sites are targets of mitogen-activated protein kinases (MAPKs) [14–16]. MAP kinase 4 (MPK4) is phosphorylated upon the invasion of *Pseudomonas syringae*, thereby releasing a coupling factor MKS1 and phosphorylate WRKY33 to target the promoter of *PHYTOALEXIN DEFICIENT 3* (PAD3) to enhance the resistance in *A. thaliana* [17]. Additionally, WRKY transcription factors also play essential roles in plant hormones signaling, including salicylic acid (SA), jasmonate acid (JA) and ethylene (ET). The WRKY transcription factors act as positive regulators in the SA signaling pathway. AtWRKY28 and AtWRKY46 are transcriptional activators of SA biosynthesis gene isochorismate synthase 1 (*ICS1*) and SA metabolism gene AVRPPHB SUSCEPTIBLE 3 (*PBS3*) [18]. In apple (*M. domestica*), MdWRKY46 is a positive regulator to enhance resistance to *Botryosphaeria dothidea* via activating the transcription of *MdPBS3.1* causing the SA signaling pathway [19]. MdWRKY15 improves resistance to *B. dothidea* via directly binding the W-box in the promoter of *MdICS1* and activating *MdICS1* transcription causing SA pathway signaling [20]. MdWRKY1 is a positive regulator of the defense response to *Alternaria blotch* or leaf spot and is induced by exogenous SA and MeJA [21]. The WRKY transcription factors act as the hub genes in the crosstalk of SA/JA and SA/ET signaling pathways. AtWRKY70 acts as a key component of the balance between signaling crosstalk promoting SA-dependent and suppressing JA-dependent responses [22]. AtWRKY33 is an important transcription factor involving in regulating the antagonistic relationship between the SA and JA disease response pathways [23,24]. The TaWRKY70 acts as a positive regulator in high-temperature seedling plant (HTSP) resistance to *Puccinia striiformis* f. sp. *tritici* (Pst.), involved in SA and ET signaling in wheat (*Triticum aestivum*) [25].

*Malus sieversii*, a tertiary relic plant, has been confirmed as the primary progenitor of all cultivated apple species [26,27]. *M. sieversii* is mainly located in central Asian, especially in Tianshan Forest of Xinjiang, China. It shows abundant genetic diversity and disease resistance features, including Fire Blight and Blue Mold, and which provide good materials for the molecular breeding of cultivated apple [28,29]. Additionally, wild apple is widely used as a quality rootstock in the agricultural production of cultivated apple, based on its excellent resistance to abiotic stress, such as drought and cold. However, suffering from *Valsa* canker disease (caused by *Valsa mali*) and invasive pest *Agrilus mali*, the population of *M. sieversii* has dramatically reduced [30,31]. Numerous research studies have reported that WRKY transcription factors play important roles to orchestrate the immunity in plants. The 25.71 Gb PacBio transcriptome and a total of 164.83 Gb Illumina database from the *M. sieversii* upon the infection of *V. mali*, provides extensive acknowledgment of the characteristic and functional identification of WRKY transcription factors in *M. sieversii* to respond to the infection of *V. mali* [32]. However, the characteristic and functional identification of WRKY family in *M. sieversii* in response to *V. mali* remains unclear. Therefore, exploring the biological function of the MsWRKY family could shed light on the stress mechanism of response to the infection of *V. mali* underlying the involvement of WRKY genes, and contribute to the improvement of *M. sieversii* germplasm and the application of gene resources to the cultivated apple.

2. Material and Methods
2.1. Plant Materials, Pathogen and Culture Conditions

The semi-annual apple seedlings (*M. domestica*, ‘M26’) were planted on the soil filling with a mixture of sterilized vermiculite and potting soil in pots. The plants were planted in
the greenhouse at 24 ± 2 °C in 16/8 h light/dark cycles. These apple seedlings were used for the transient transformation assay. The tissue-cultured apple seedlings of *M. sieversii* were cultured in a climate chamber with a day/night cycle of 16/8 h and at 24 ± 2 °C at a light intensity of 25 µmol m⁻² s⁻¹. The fungal pathogen *V. mali* isolate EGI 1 was isolated from the *M. sieversii* in the Tianshan Wild Fruit Forest area of Xinjiang, China, and identified by Liu et al. [31]. The *V. mali* isolate EGI 1 was grown on the (potato dextrose agar) PDA plate at 25 °C for three days.

### 2.2. Identification of the WRKY Family Proteins in *M. Sieversii*

The transcriptome database of *M. sieversii* upon *V. mali* infection was derived from the laboratory of Professor Daoyuan Zhang, Xinjiang Institute of Ecology and Geography Chinese Academy of Sciences, which was published with the project accession number PRJNA687214 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA687214?reviewer=dve4c2uoschi4ag7dnbunbr44, accessed on 22 December 2020) [32]. For the sampling of the PacBio and Illumina sequencing, barks of twigs in *M. sieversii* near the canker were separately harvested at the time points of 0, 1, 2, and 5 dpi and each sample contained three biological replicates. These uninfected bark samples at 0 dpi time point were collected for RNA extraction as controls. The hidden Markov models (PF03106) of WRKY protein was downloaded from pfam database (http://pfam.xfam.org/, accessed on 1 January 2019). Hmmmscan program of HMMER package was used to search the potential WRKY protein form the transcriptome database of *M. sieversii* with parameters as “E cut-off 1e”. The sequences were confirmed using the NCBI-CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml, accessed on 9 May 2019) and SMART databases (http://smart.embl-heidelberg.de/, accessed on 2 August 2019) [33]. After removing the redundant and incorrect predicted-protein sequences, the MsWRKY proteins were validated finally.

### 2.3. Protein Property Identification of MsWRKY Proteins

We used the ExPASy Protparam (https://web.expasy.org/protparam/, accessed on) to examine protein properties of the MsWRKYs, including molecular weight (MW), hydrophobicity, instability index and isoelectric point (PI). Phosphorylation sites of MsWRKY proteins were predicted by KinasePhos (http://kinasephos.mblv.nctu.edu.tw/predict.php, accessed on 16 May 2019). Membrane binding domain was analyzed by TMHMM Server v 2.0 (http://www.cbs.dtu.dk/services/TMHMM, accessed on 9 May 2019) and Phobius (https://www.ebi.alf.uk/Tools/pfa/phobius/, accessed on 19 March 2019). cNLS Mapper (http://nls-mapper.iba.keio.aV.jp/cgiin/NLS_Mapper_form.cgi, accessed on 26 February 2019) was used to predict the nuclear localization signal (NLS). The putative subcellular locations were analyzed by WoLF PSORT (https://wolfpsort.hgV.jp/, accessed on 26 February 2019).

### 2.4. Classification and Phylogenetic Analysis of MsWRKYs

Multiple alignments were conducted by DNAMAN Version 9.0 with an identity protein weight matrix to obtain detailed information on the structure of MsWRKYs. Weblogo (https://weblogo.berkeley.edu/logo.cgi, accessed on 8 May 2019) was used to generate the conversed motif logo of MsWRKYs. MEME suite 5.1.1 (http://meme-suite.org, accessed on 8 May 2019) was employed to analyze the possible conversed motifs in MsWRKY protein sequences setting a maximum value of motifs to 15. The motif patterns were redrawn by TBtools software (http://cj-chen.github.io/tbtools/, accessed on 8 May 2019).

The AtWRKY protein sequences were downloaded from TAIR (http://www.arabidopsis.org/, accessed on 29 April 2019), and were used to classified the WRKY protein in *M. sieversii*. Phylogenetic analysis was performed with a neighbor-joining (NJ) method, with 1000 bootstrap replicates and Poisson Correction model in MEGA 7.0. The phylogenetic tree was modified by Figtree v1.4.4.
2.5. RNA Extraction and qRT-PCR Analysis

The total RNAs used for the constructions of Illumina and PacBio sequencing libraries were used to implement the qRT-PCR assay, which was for validation of the consistency of RNA-seq analysis. Apple leaves in the transient transformation assay were collected and full ground with liquid nitrogen. Total RNAs were isolated using Plant RNA Kit (Omega, No. R6827, China). The first-strand cDNA was synthesized with 1 µg of total RNA from each sample using PrimeScript RT reagent Kit with gDNA Eraser (Takara, No. RR047Q, China). Gene-specific primers used for the validation of expression levels of MsWRKYs (Table S3), were designed using primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/, accessed on 29 November 2018) and synthesized by Sangon Biotech (Shanghai, China). Real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) was performed with TB Green Premix Ex Taq II kit (Takara, No. RR820A, China) on CFX96 Real-Time PCR Detection System (Bio-Rad, CFX96, American). The reaction system contained 12.5 µL TB Green Premix Ex Taq (2×) buffer, 1 µL each of the primers (10 µM) and 2 µL template and 8.5 µL ddH₂O. The thermal profile for qRT-PCR was as follows: preheating at 95 °C for 30 s; cycling stage: 95 °C for 5 s, 60 °C for 30 s, 40 cycles; melt curve analysis was produced to confirm the specificity of amplification. The relative transcript abundances were analyzed with the 2⁻∆∆CT method [34]. The EF-1α (Elongation factor 1-α) gene was used as an internal reference gene [35]. Each sample contained three biological replicates and each biological replicate contained three technical replicates. Statistical analysis was performed by one-way ANOVA method, Tukey’s HSD test using SPSS v.18 software. The heatmap of each group of MsWRKYs was generated by TBtools (http://cj-chen.github.io/tbtools/, accessed on 5 May 2019) based on the FPKM values of the Illumina transcriptome database.

2.6. Transient Expression Assay

To transiently express the selected MsWRKY transcripts, pBI121-MsWRKYs-GFP over-expression vectors were constructed using an in-fusion HD cloning kit (Code No. 639648). The method of transient expression was performed followed as previously described with modifications [36]. The A. tumefaciens in 20% glycerol stock at −80 °C was recovered on a Luria–Bertani (LB) agar plate for 3-day incubation at 28 °C. A single colony from the plate was picked up to inoculate in 5 mL LB medium containing kanamycin (50 µg/mL) and rifampicin (50 µg/mL) for shaking (220 rpm) at 28 °C for 20–24 h. Moreover, 1 mL of the overnight grown bacterial solution was taken into the sterile 100 mL LB medium with antibiotics and continued to grow for 6–8 h until the OD₆₅₀ of 0.6. For pre-induction, the pellet of A. tumefaciens was suspended with AB-MES solution (17.2 mM K₂HPO₄, 8.3 mM NaH₂PO₄, 18.7 mM NH₄Cl, 2 mM KCl, 1.25 mM MgSO₄, 100 µM CaCl₂, 10 µM FeSO₄, 50 mM MES, 2% glucose (w/v), pH 5.5) and acetosyringone (AS, 200 µM) to OD₆₅₀ of 0.2 without antibiotics, then shaken (220 rpm) at 28 °C for 12–16 h. To perform the transient transfection procedure, the induced A. tumefaciens were collected and resuspended with the solution AB-MES and MS medium (v/v: 1:1) containing the AS (100 µM). The bacterial solution used for transient transfection was adjusted to OD₆₅₀ of 0.8. The detached leaves were washed with ddH₂O and disinfected the surface with alcohol cotton. The leaves were immersed in the transfection solution containing A. tumefaciens in the dark at 28 °C for 1 h. Subsequently, these washed leaves with ddH₂O were placed in sterile dishes and preserved moisture. After a one-day cocultivation in the dark at 28 °C, the transiently expressed leaves were prepared for inoculations of V. mali.

2.7. Infection Assays with Transient Expression Leaves

The procedures for the infection assay used the method optimized by Liu et al. [31]. The mycelial plugs (diameter: 5 mm) were excised aseptically from the edge of the 3-days-cultured isolate EGI1 on PDA medium. After the 2-day transient expression, these leaves were wounded by a fabric pattern wheel and inoculated with these mycelial plugs. Leaves inoculated with ddH₂O acted as controls. The mocks were that leaves were inoculated with
For the hormone treatment, apple seedlings (M. domestica ‘M26’) were planted in the greenhouse condition, which was 25 ± 2 °C and a relative humidity of 50 ± 10%. Leaves were detached to use to inoculate with V. mali and spray in the solutions, which were SA solution (5 mM, dissolved in sterile ddH₂O), MeJA solution (1 mM, dissolved in sterile ddH₂O) and ACC solution (1 mM, dissolved in sterile ddH₂O) [37,38]. After the inoculation of V. mali and hormone treatments, the leaf samples were immediately frozen in liquid nitrogen at 0 h, 1 h, 3 h, 6 h, 1 d, 2 d, 5 d timepoints, and stored at −80 °C for further use. The hormone treatment was implemented in a growth chamber at 24 °C in 16/8 h light/dark cycles and 60% relative humidity.

3. Results

A total of 112 unique MsWRKY transcripts encoding the conserved WRKY domain were identified from the genome of M. domestica, named as MsWRKY1 to MsWRKY112. The MsWRKY106 had the largest amino acid sequence of 995 aa (MW: 11.32 kDa), while the MsWRKY55 had the smallest amino acid sequence of 80 aa (MW: 9.28 kDa). The isoelectric points (PIs) ranged from 4.81 (MsWRKY83)–9.99 (MsWRKY78). The highly conserved domain WRKYGQK was presented in 110 MsWRKY proteins, whereas the remaining two members contained WRKYGKK and WRKYGMK domains (Table 1). The biochemical properties of MsWRKY proteins were predicted, including the grand average of hydropathicity (GRAVY), instability index, subcellular location, nuclear location signal (NLS) and phosphorylation site. One hundred and eight MsWRKY proteins were hydrophilic protein (GRAVY < −0.5); the remaining four MsWRKY proteins were hydrophobic (GRAVY > −0.5). The instability indexes of 105 MsWRKY proteins were more than 40, which showed that they were unstable proteins. The predicted subcellular location showed that MsWRKY proteins localized in the nucleus (75), chloroplast (27) and cytoplasm (10). It was inferred that the MsWRKY transcription factors outside of the nucleus might play transcriptional regulatory functions through the process of transmembrane transfer. NLS prediction analysis showed that 27 of 112 MsWRKY proteins were without NLSs. The phosphorylation site prediction analyzed by KinasePhos showed that MsWRKY proteins had abundant phosphorylation sites, including predominant serine (S), threonine (T) and tyrosine (Y) sites. There were 43 phosphorylation sites in the MsWRKY29 protein, while there were 2 phosphorylation sites in MsWRKY55 and MsWRKY60 separately (Table S1).
Table 1. List of MsWRKY transcripts identified in *M. sieversii*.

| Name | ID          | Group | ORF (aa) | MW    | PI   | Conserved Heptapeptide | Zinc-Finger Type | Domain Number | ID            | Orthology in *A. thaliana* |
|------|-------------|-------|----------|-------|------|------------------------|------------------|---------------|---------------|--------------------------|
| MsWRKY1 | MD02G1007900 | I     | 473      | 52,029.15 | 8.82 | WRKYGQK               | C2H2             | 2             | AT4G30935.1 | WRKY32                   |
| MsWRKY2 | MD03G1044400 | I     | 732      | 79,512.24 | 5.96 | WRKYGQK               | C2H2             | 2             | AT5G56270.1 | WRKY2                    |
| MsWRKY3 | MD03G1188900 | I     | 584      | 63,851.08 | 5.97 | WRKYGQK               | C2H2             | 2             | AT4G26640.2 | WRKY20                   |
| MsWRKY4 | MD04G1113100 | I     | 470      | 51,522.09 | 8.93 | WRKYGQK               | C2H2             | 2             | AT2G37260.1 | WRKY40                   |
| MsWRKY5 | MD04G1167700 | I     | 520      | 57,708.98 | 7.1  | WRKYGQK               | C2H2             | 2             | AT2G38470.1 | WRKY33                   |
| MsWRKY6 | MD04G1244700 | I     | 717      | 78,545.69 | 5.99 | WRKYGQK               | C2H2             | 2             | AT5G56270.1 | WRKY2                    |
| MsWRKY7 | MD06G1155200 | I     | 924      | 102,781.88 | 5.33 | WRKYGQK               | C2H2             | 2             | AT2G38470.1 | WRKY33                   |
| MsWRKY8 | MD09G1056600 | I     | 528      | 57,615.66 | 8.62 | WRKYGQK               | C2H2             | 2             | AT2G03340.1 | WRKY3                    |
| MsWRKY9 | MD09G1121600 | I     | 484      | 53,154.16 | 5.93 | WRKYGQK               | C2H2             | 2             | AT2G04880.1 | WRKY1                    |
| MsWRKY10 | MD11G1059400 | I     | 572      | 62,895.54 | 6.77 | WRKYGQK               | C2H2             | 2             | AT2G38470.1 | WRKY33                   |
| MsWRKY11 | MD11G1205000 | I     | 588      | 63,957.32 | 5.86 | WRKYGQK               | C2H2             | 2             | AT4G266640.2 | WRKY20                   |
| MsWRKY12 | MD12G1128800 | I     | 470      | 51,455.11 | 9    | WRKYGQK               | C2H2             | 2             | AT2G37260.1 | WRKY40                   |
| MsWRKY13 | MD12G1810000 | I     | 512      | 56,720.99 | 6.73 | WRKYGQK               | C2H2             | 2             | AT2G38470.1 | WRKY33                   |
| MsWRKY14 | MD12G1260600 | I     | 718      | 78,957.76 | 6.4  | WRKYGQK               | C2H2             | 2             | AT5G56270.1 | WRKY2                    |
| MsWRKY15 | MD13G1067600 | I     | 526      | 57,344.94 | 7.26 | WRKYGQK               | C2H2             | 2             | AT2G03340.1 | WRKY3                    |
| MsWRKY16 | MD03G1057400 | I     | 571      | 62,545.25 | 7.02 | WRKYGQK               | C2H2             | 2             | AT2G38470.1 | WRKY33                   |
| MsWRKY17 | MD15G1321000 | I     | 434      | 48,117.41 | 7.27 | WRKYGQK               | C2H2             | 2             | AT4G30935.1 | WRKY32                   |
| MsWRKY18 | MD16G1066500 | I     | 528      | 57,197.02 | 8.38 | WRKYGQK               | C2H2             | 2             | AT2G03340.1 | WRKY3                    |
| MsWRKY19 | MD17G1054100 | I     | 530      | 57,383.57 | 8.39 | WRKYGQK               | C2H2             | 2             | AT2G03340.1 | WRKY3                    |
| MsWRKY20 | MD17G1126000 | I     | 471      | 51,661.31 | 6.64 | WRKYGQK               | C2H2             | 2             | AT2G04880.1 | WRKY1                    |
| MsWRKY21 | MD00G1143500 | II-a  | 334      | 36,975.49 | 7    | WRKYGQK               | C2H2             | 1             | AT1G80840.1 | WRKY40                   |
| MsWRKY22 | MD00G1143600 | II-a  | 278      | 31,279.14 | 8.85 | WRKYGQK               | C2H2             | 1             | AT1G80840.1 | WRKY40                   |
| MsWRKY23 | MD09G1224500 | II-a  | 320      | 35,309.59 | 7.59 | WRKYGQK               | C2H2             | 1             | AT1G80840.1 | WRKY40                   |
| MsWRKY24 | MD15G1039500 | II-a  | 302      | 33,600.37 | 7.1  | WRKYGQK               | C2H2             | 1             | AT4G31800.1 | WRKY18                   |
| MsWRKY25 | MD15G1039600 | II-a  | 286      | 32,088.00 | 8.15 | WRKYGQK               | C2H2             | 1             | AT1G80840.1 | WRKY40                   |
| MsWRKY26 | MD17G1223100 | II-a  | 321      | 35,670.75 | 8.2  | WRKYGQK               | C2H2             | 1             | AT1G80840.1 | WRKY40                   |
| MsWRKY27 | MD03G1979600 | II-b  | 538      | 58,927.85 | 6.42 | WRKYGQK               | C2H2             | 1             | AT1G62300.1 | WRKY6                    |
| MsWRKY28 | MD05G1349800 | II-b  | 606      | 65,345.2  | 7.67 | WRKYGQK               | C2H2             | 1             | AT1G62300.1 | WRKY6                    |
| MsWRKY29 | MD06G1189100 | II-b  | 683      | 73,465.97 | 7.23 | WRKYGQK               | C2H2             | 1             | AT1G18860.1 | WRKY61                   |
| MsWRKY30 | MD08G1227200 | II-b  | 570      | 62,311.55 | 5.35 | WRKYGQK               | C2H2             | 1             | AT1G68150.1 | WRKY9                    |
| MsWRKY31 | MD09G1048300 | II-b  | 470      | 51,182.97 | 7.61 | WRKYGQK               | C2H2             | 1             | AT5G15130.1 | WRKY72                   |
| MsWRKY32 | MD09G1111200 | II-b  | 625      | 68,204.54 | 6.33 | WRKYGQK               | C2H2             | 1             | AT1G62300.1 | WRKY6                    |
| MsWRKY33 | MD10G1324500 | II-b  | 611      | 65,654.44 | 6.61 | WRKYGQK               | C2H2             | 1             | AT1G62300.1 | WRKY6                    |
| MsWRKY34 | MD11G1213500 | II-b  | 541      | 58,972.76 | 6.45 | WRKYGQK               | C2H2             | 1             | AT1G62300.1 | WRKY6                    |
| MsWRKY35 | MD13G1077900 | II-b  | 571      | 61,852.99 | 5.95 | WRKYGQK               | C2H2             | 1             | AT1G18860.1 | WRKY61                   |
Table 1. Cont.

| Name       | ID          | Group | ORF (aa) | MW   | PI    | Conserved Heptapeptide | Zinc-Finger Type | Domain Number | ID          | Orthology in A. thaliana |
|------------|-------------|-------|----------|------|-------|------------------------|------------------|---------------|-------------|--------------------------|
| MsWRKY36   | MD14G1196100| II-b  | 584      | 63,307.85 | 6.48 | WRKYGQK                | C2H2             | 1             | AT5G15130.1 | WRKY72                   |
| MsWRKY37   | MD15G149600 | II-b  | 582      | 64,167.53 | 5.1  | WRKYGQK                | C2H2             | 1             | AT1G68150.1 | WRKY9                    |
| MsWRKY38   | MD16G1077700| II-b  | 587      | 64,141.44 | 6.43 | WRKYGQK                | C2H2             | 1             | AT5G15130.1 | WRKY72                   |
| MsWRKY39   | MD17G1048400| II-b  | 445      | 49,960.64 | 7.59 | WRKYGQK                | C2H2             | 1             | AT5G15130.1 | WRKY72                   |
| MsWRKY40   | MD17G1099000| II-b  | 645      | 69,960.43 | 6.42 | WRKYGQK                | C2H2             | 1             | AT1G62300.1 | WRKY6                    |
| MsWRKY41   | MD01G1033500| II-c  | 270      | 30,431.96 | 8.93 | WRKYGQK                | C2H2             | 1             | AT4G39410.1 | WRKY13                   |
| MsWRKY42   | MD01G1071600| II-c  | 218      | 24,560.79 | 9.36 | WRKYGQK                | C2H2             | 1             | AT1G64000.1 | WRKY56                   |
| MsWRKY43   | MD01G1239000| II-c  | 208      | 23,620.15 | 8.91 | WRKYGQK                | C2H2             | 1             | AT5G41570.1 | WRKY24                   |
| MsWRKY44   | MD04G1112800| II-c  | 146      | 16,658.26 | 9.6  | WRKYGQK                | C2H2             | 1             | AT2G37260.1 | WRKY44                   |
| MsWRKY45   | MD04G131000 | II-c  | 297      | 33,413.28 | 5.72 | WRKYGQK                | C2H2             | 1             | AT5G43290.1 | WRKY49                   |
| MsWRKY46   | MD05G1290300| II-c  | 319      | 35,119.81 | 6.6  | WRKYGQK                | C2H2             | 1             | AT1G29960.1 | WRKY71                   |
| MsWRKY47   | MD06G1138500| II-c  | 150      | 17,147.32 | 9.56 | WRKYGQK                | C2H2             | 1             | AT5G13080.1 | WRKY75                   |
| MsWRKY48   | MD07G1110400| II-c  | 236      | 26,784.29 | 8.17 | WRKYGQK                | C2H2             | 1             | AT2G44745.1 | WRKY12                   |
| MsWRKY49   | MD07G1131000| II-c  | 222      | 24,930.24 | 9.36 | WRKYGQK                | C2H2             | 1             | AT1G64000.1 | WRKY56                   |
| MsWRKY50   | MD08G1067700| II-c  | 161      | 18,241.06 | 8.53 | WRKYGQK                | C2H2             | 1             | AT5G26170.1 | WRKY50                   |
| MsWRKY51   | MD09G1008800| II-c  | 221      | 25,441.17 | 9.12 | WRKYGQK                | C2H2             | 1             | AT5G13080.1 | WRKY75                   |
| MsWRKY52   | MD09G1507000| II-c  | 369      | 41,364.41 | 7.75 | WRKYGQK                | C2H2             | 1             | AT1G29960.1 | WRKY71                   |
| MsWRKY53   | MD09G1285400| II-c  | 350      | 38,393.32 | 5.85 | WRKYGQK                | C2H2             | 1             | AT2G47260.1 | WRKY23                   |
| MsWRKY54   | MD10G1266400| II-c  | 327      | 35,938.73 | 6.46 | WRKYGQK                | C2H2             | 1             | AT4G18170.1 | WRKY28                   |
| MsWRKY55   | MD12G1199000| II-c  | 80       | 9282.47   | 9.85 | WRKYGQK                | C2H2             | 1             | AT2G03340.1 | WRKY3                    |
| MsWRKY56   | MD12G1441000| II-c  | 297      | 33,084.78 | 5.94 | WRKYGQK                | C2H2             | 1             | AT5G43290.1 | WRKY49                   |
| MsWRKY57   | MD13G1064700| II-c  | 346      | 37,734.7  | 6.59 | WRKYGQK                | C2H2             | 1             | AT1G69310.1 | WRKY57                   |
| MsWRKY58   | MD13G1221000| II-c  | 190      | 21,714.42 | 9.74 | WRKYGQK                | C2H2             | 1             | AT5G13080.1 | WRKY75                   |
| MsWRKY59   | MD13G1507000| II-c  | 385      | 42,703.88 | 6.11 | WRKYGQK                | C2H2             | 1             | AT5G49520.1 | WRKY48                   |
| MsWRKY60   | MD14G1154500| II-c  | 148      | 17,090.15 | 9.59 | WRKYGQK                | C2H2             | 1             | AT3G01970.1 | WRKY45                   |
| MsWRKY61   | MD15G1054000| II-c  | 161      | 18,178.8  | 5.65 | WRKYGQK                | C2H2             | 1             | AT5G26170.1 | WRKY50                   |
| MsWRKY62   | MD15G1331300| II-c  | 199      | 22,265.22 | 5.86 | WRKYGQK                | C2H2             | 1             | AT5G64810.1 | WRKY51                   |
| MsWRKY63   | MD15G1337100| II-c  | 271      | 30,236.73 | 8.93 | WRKYGQK                | C2H2             | 1             | AT4G39410.1 | WRKY13                   |
| MsWRKY64   | MD16G1063200| II-c  | 324      | 35,540.18 | 6.6  | WRKYGQK                | C2H2             | 1             | AT1G69310.1 | WRKY57                   |
| MsWRKY65   | MD16G1122400| II-c  | 161      | 18,041.02 | 9.08 | WRKYGQK                | C2H2             | 1             | AT5G13080.1 | WRKY75                   |
| MsWRKY66   | MD16G1151000| II-c  | 371      | 41,144.21 | 5.6  | WRKYGQK                | C2H2             | 1             | AT5G49520.1 | WRKY48                   |
| MsWRKY67   | MD17G1001500| II-c  | 223      | 25,599.42 | 9.14 | WRKYGQK                | C2H2             | 1             | AT5G13080.1 | WRKY75                   |
| MsWRKY68   | MD17G1138100| II-c  | 365      | 41,251.08 | 6.76 | WRKYGQK                | C2H2             | 1             | AT4G18170.1 | WRKY28                   |
| MsWRKY69   | MD17G1278100| II-c  | 346      | 38,166.27 | 6.27 | WRKYGQK                | C2H2             | 1             | AT2G47260.1 | WRKY23                   |
| MsWRKY70   | MD02G1177500| II-d  | 330      | 35,972.67 | 9.65 | WRKYGQK                | C2H2             | 1             | AT4G24240.1 | WRKY7                    |
| Name ID    | Group | ORF (aa) | MW     | PI    | Conserved Heptapeptide | Zinc-Finger Type | Domain Number | ID          | Orthology in A. thaliana |
|-----------|-------|----------|--------|-------|------------------------|------------------|---------------|------------|--------------------------|
| MsWRKY71 MD04G1226400 | II-d  | 325      | 36,589.02 | 9.77  | WRKYGQK                | C2H2             | 1            | AT5G28650.1 | WRKY74                  |
| MsWRKY72 MD05G1204400 | II-d  | 354      | 39,961.26  | 9.7   | WRKYGQK                | C2H2             | 1            | AT2G30590.1 | WRKY21                  |
| MsWRKY73 MD06G1062800 | II-d  | 318      | 35,883.94  | 9.58  | WRKYGQK                | C2H2             | 1            | AT2G30590.1 | WRKY21                  |
| MsWRKY74 MD08G1094900 | II-d  | 356      | 38,637.1   | 9.41  | WRKYGQK                | C2H2             | 1            | AT4G24240.1 | WRKY7                   |
| MsWRKY75 MD08G1127200 | II-d  | 341      | 36,964.66  | 9.32  | WRKYGQK                | C2H2             | 1            | AT4G31550.2 | WRKY11                  |
| MsWRKY76 MD10G1191400 | II-d  | 355      | 39,953.34  | 9.68  | WRKYGQK                | C2H2             | 1            | AT2G30590.1 | WRKY21                  |
| MsWRKY77 MD12G1243400 | II-d  | 313      | 35,302.81  | 9.96  | WRKYGQK                | C2H2             | 1            | AT5G28650.1 | WRKY74                  |
| MsWRKY78 MD13G1239100 | II-d  | 281      | 30,732.72  | 9.99  | WRKYGQK                | C2H2             | 1            | AT4G31550.2 | WRKY11                  |
| MsWRKY79 MD15G1078200 | II-d  | 342      | 37,212.39  | 9.26  | WRKYGQK                | C2H2             | 1            | AT4G24240.1 | WRKY7                   |
| MsWRKY80 MD15G1106600 | II-d  | 338      | 36,702.39  | 9.46  | WRKYGQK                | C2H2             | 1            | AT4G31550.1 | WRKY11                  |
| MsWRKY81 MD15G1287300 | II-d  | 331      | 36,240.86  | 9.54  | WRKYGQK                | C2H2             | 1            | AT2G23320.1 | WRKY15                  |
| MsWRKY82 MD16G1244300 | II-d  | 284      | 30,922.13  | 9.87  | WRKYGQK                | C2H2             | 1            | AT5G15130.1 | WRKY72                  |
| MsWRKY83 MD00G1140800 | II-e  | 268      | 30,828.52  | 4.81  | WRKYGQK                | C2H2             | 1            | AT2G34830.2 | WRKY35                  |
| MsWRKY84 MD03G1292900 | II-e  | 260      | 30,342.43  | 5.17  | WRKYGQK                | C2H2             | 1            | AT2G34830.2 | WRKY35                  |
| MsWRKY85 MD05G1295700 | II-e  | 273      | 30,201.82  | 5.8   | WRKYGQK                | C2H2             | 1            | AT1G29280.1 | WRKY65                  |
| MsWRKY86 MD09G1235100 | II-e  | 250      | 27,398.47  | 5.47  | WRKYGQK                | C2H2             | 1            | AT3G58710.1 | WRKY69                  |
| MsWRKY87 MD10G1275800 | II-e  | 266      | 29,502.25  | 5.1   | WRKYGQK                | C2H2             | 1            | AT1G29280.1 | WRKY65                  |
| MsWRKY88 MD01G1071300 | II-e  | 349      | 37,654.85  | 8.16  | WRKYGQK                | C2H2             | 1            | AT4G01250.1 | WRKY22                  |
| MsWRKY89 MD01G1210200 | II-e  | 470      | 51,759.28  | 5.08  | WRKYGQK                | C2H2             | 1            | AT5G52830.1 | WRKY27                  |
| MsWRKY90 MD05G1265200 | II-e  | 492      | 53,777.42  | 5.85  | WRKYGQK                | C2H2             | 1            | AT1G30650.1 | WRKY14                  |
| MsWRKY91 MD06G1091200 | II-e  | 295      | 33,549.62  | 5.49  | WRKYGQK                | C2H2             | 1            | AT4G23350.1 | WRKY21                  |
| MsWRKY92 MD07G1131400 | II-e  | 348      | 37,658.79  | 6.75  | WRKYGQK                | C2H2             | 1            | AT4G01250.1 | WRKY22                  |
| MsWRKY93 MD07G1280300 | II-e  | 278      | 31,066.3   | 7.1   | WRKYGQK                | C2H2             | 1            | AT5G52830.1 | WRKY27                  |
| MsWRKY94 MD10G1243000 | II-e  | 493      | 53,237.86  | 6.05  | WRKYGQK                | C2H2             | 1            | AT1G30650.1 | WRKY14                  |
| MsWRKY95 MD14G1112200 | II-e  | 315      | 35,646.87  | 5.05  | WRKYGQK                | C2H2             | 1            | AT4G23350.1 | WRKY29                  |
| MsWRKY96 MD01G1078000 | III   | 353      | 39,173.67  | 5.3   | WRKYGQK                | C2HC             | 1            | AT4G23810.1 | WRKY53                  |
| MsWRKY97 MD01G1168500 | III   | 228      | 25,662.99  | 7.72  | WRKYGQK                | C2HC             | 1            | AT2G40704.1 | WRKY55                  |
| MsWRKY98 MD01G1168600 | III   | 303      | 34,226.04  | 5.71  | WRKYGQK                | C2HC             | 1            | AT3G56400.1 | WRKY70                  |
| MsWRKY99 MD01G1215300 | III   | 347      | 38,710.64  | 5.93  | WRKYGQK                | C2HC             | 1            | AT4G23810.1 | WRKY53                  |
| MsWRKY100 MD04G1175500 | III   | 333      | 37,191.56  | 5.38  | WRKYGQK                | C2HC             | 1            | AT5G22570.1 | WRKY38                  |
| MsWRKY101 MD04G1175600 | III   | 344      | 38,249.03  | 6.01  | WRKYGQK                | C2HC             | 1            | AT5G22570.1 | WRKY38                  |
| MsWRKY102 MD06G1041000 | III   | 351      | 39,416.65  | 5.72  | WRKYGQK                | C2HC             | 1            | AT4G11070.1 | WRKY41                  |
| MsWRKY103 MD07G1146900 | III   | 356      | 39,731.16  | 5.48  | WRKYGQK                | C2HC             | 1            | AT4G23810.1 | WRKY53                  |
| MsWRKY104 MD07G1234600 | III   | 342      | 37,441.69  | 5.48  | WRKYGQK                | C2HC             | 1            | AT2G40704.1 | WRKY55                  |
| MsWRKY105 MD07G1234700 | III   | 302      | 33,972.88  | 6.38  | WRKYGQK                | C2HC             | 1            | AT3G56400.1 | WRKY70                  |
| Name      | ID          | Group | ORF (aa) | MW        | PI   | Conserved Heptapeptide | Zinc-Finger Type | Domain Number | ID      | Orthology in A. thaliana |
|-----------|-------------|-------|----------|-----------|------|------------------------|------------------|---------------|---------|-----------------------|
| MsWRKY106 | MD07G1261100| III   | 995      | 113,186.85| 8.64 | WRKYGQK                | C2HC             | 1             | /       | /                     |
| MsWRKY107 | MD07G1285200| III   | 342      | 38,048.86 | 5.5  | WRKYGQK                | C2HC             | 1             | AT4G11070.1 | WRKY41                |
| MsWRKY108 | MD12G1189200| III   | 339      | 37,732.39 | 6.41 | WRKYGQK                | C2HC             | 1             | AT3G56400.1 | WRKY70                |
| MsWRKY109 | MD12G1189600| III   | 369      | 40,597.61 | 5.69 | WRKYGQK                | C2HC             | 1             | AT3G56400.1 | WRKY70                |
| MsWRKY110 | MD12G1189700| III   | 332      | 37,182.61 | 5.77 | WRKYGQK                | C2HC             | 1             | AT3G56400.1 | WRKY70                |
| MsWRKY111 | MD12G1189900| III   | 237      | 27,258.08 | 8.26 | WRKYGQK                | C2HC             | 1             | AT2G46400.1 | WRKY46                |
| MsWRKY112 | MD14G1123000| III   | 355      | 39,872.02 | 5.66 | WRKYGQK                | C2HC             | 1             | AT4G11070.1 | WRKY41                |
3.2. Phylogenetic Analysis and Classification of MsWRKY Proteins

Based on the number of WRKY domains and the zinc finger type, the putative WRKY proteins could be grouped I, II (a–e) and III [7]. Multiple alignment results showed that Group I (MsWRKY1–20) had two WRKYGQK domains and zinc-finger motifs of C2H2 type. Group II (MsWRKY21–95) had one WRKYGQK domain and two zinc-finger motifs of C2H2 type except for MsWRKY50/MsWRKY61 (WRKYGKK domain) and MsWRKY62 (WRKYGMK domain). Aligned with AtWRKY protein sequences, Group II of MsWRKY proteins were classified into the five subgroups (II-a, II-b, II-c, II-d, II-e). Sixteen members of Group III (MsWRKY96–112) had one WRKYGQK domain and a zinc-finger motif of C2HC (Table 1). Additionally, the zinc-finger motifs of MsWRKY proteins were relatively conserved. Group I N, C and group II-c had the zinc-finger motif of CX4CX23HXC, except for MsWRKY6 (CX4CX24HXH) and MsWRKY14 (CX4CX23HXH). Group II-a, II-b, II-d and II-e had the zinc-finger motif of CX5CX23HXH. Excluding MsWRKY100 and MsWRKY101 (CX5CX23HXC), the Group III MsWRKYs had zinc-finger of CX7CX23HXC (Table 1). To determine the phylogenetic relationship between all 112 MsWRKYs, the neighbor-joining (NJ) phylogenetic tree was constructed with 112 MsWRKYs, 65 AtWRKYs and 113 Md-WRKYs proteins (Figure 1). The NJ phylogenetic tree of MsWRKY full-length proteins was clustered into three groups (I, II, and III). The Group II diverged into three clades, of which were clade IIa/IIb, clade IIc/IIe and clade IIV. It indicated that close evolutionary relationships were between subgroup IIa and IIb; also, IId and Ile.

3.3. Conserved Motif Composition Analysis of MsWRKY Proteins

The conserved motifs of 112 MsWRKY proteins were predicted using MEME suite 5.1.1 and Weblogo. The results showed that the fifteen conserved motifs in MsWRKY proteins. The different groups of MsWRKYs had characteristic consistent conserved motifs. The WRKY domains of group I-N were constituted of motif 4 and 6, while group I-C was constituted with motif 1, 3 and 2. Additionally, the N-terminal of I-C and IIc was motif 5. Additionally, except for the group I-N and III, the N-terminal of the WRKY domain consisted of characteristic motifs. The N-terminal of WRKY domain of group (I-C and IIc), group (IIa and IIb), and group (IId and Ile) consisted of the motifs 5, 12 and 11, respectively. For the N-terminal of MsWRKY proteins, there were different conservative motifs. The N-terminal of MsWRKY proteins of group I, group (IIa and IIb), group IId and group III consisted of the motifs 9, 7, 13 and (13 + 14), respectively (Figure 2).

3.4. Transmembrane Domain (TMD) and Conserved Domain in MsWRKY Proteins

Through TMHMM and SMART prediction analysis, four MsWRKY proteins with transmembrane domains were identified (Table S1). The TMDs of MsWRKY7, MsWRKY97 and MsWRKY111 were in the N-terminus, while the TMD of MsWRKY44 protein was in the C-terminus. Additionally, six specific conserved domains were identified in certain MsWRKY proteins, including RPW8, rad50, UPF0242, nuc hydro, Plant zn clust and PLN03210 domains using NCBI CDD database. There were two RPW8 domains in the N-terminal of the WRKY domain of the MsWRKY7 protein. A rad50 and UPF0242 domains were separately identified in the N-terminal of the WRKY domain of MsWRKY28 and MsWRKY29 proteins. A nuc hydro domain was founded in the C-terminal of the WRKY domain of the MsWRKY31 protein. The MsWRKY70–82 proteins each had one plant zn clust in the N-terminal of the WRKY domain (Figure 3).
Figure 1. Phylogenetic analysis of the WRKY proteins among *M. sieversii* (red), *M. domestica* (blue) and *A. thaliana* (black). The phylogenetic tree was constructed using the Neighbor-Joining method. The bootstrap test was performed with 1000 replicates. The WRKY proteins of *M. sieversii*, *M. domestica* and *A. thaliana* were marked in red, blue and black.
3.4. Transmembrane Domain (TMD) and Conserved Domain in MsWRKY Proteins

Through TMHMM and SMART prediction analysis, four MsWRKY proteins with transmembrane domains were identified (Table S1). The TMDs of MsWRKY7, MsWRKY97 and MsWRKY111 were in the N-terminus, while the TMD of MsWRKY44 protein was in the C-terminus. Additionally, six specific conserved domains were identified in certain MsWRKY proteins, including RPW8, rad50, UPF0242, nuc hydro, Plant zn clust and PLN03210 domains using NCBI CDD database. There were two RPW8 domains in the N-terminal of the WRKY domain of the MsWRKY7 protein. A rad50 and UPF0242 domains were separately identified in the N-terminal of the WRKY domain of MsWRKY28 and MsWRKY29 proteins. A nuc hydro domain was founded in the C-terminal of the WRKY domain of the MsWRKY31 protein. The MsWRKY7–0–82 proteins each had one plant zn clust in the N-terminal of the WRKY domain (Figure 3).

3.5. Expression Patterns of MsWRKY Transcripts upon the V. mali Infection

To confirm whether MsWRKYs are involved in response to the infection of V. mali, expression patterns of the 112 MsWRKY transcripts were analyzed upon the infection of V. mali. (Figure 4A, Table S2). The expression levels were represented using FPKM values at 0, 1, 2, 5 dpi. The transcripts were considered to be differentially expressed depending on the cut-off that was $|\log_2$(Fold change)$| \geq 1$ with the Q-value < 0.05. Thirty-three MsWRKY transcripts were significantly differentially expressed during the response stage, indicating that these MsWRKY transcripts may play important roles in response to V. mali. The differentially expressed MsWRKY transcripts were clustered into three clusters based on expression patterns. Cluster 1 contained 10 MsWRKY transcripts that were all up-regulated at 1, 2 and 5 dpi, including MsWRKY6, MsWRKY27, MsWRKY29, MsWRKY30, MsWRKY34, MsWRKY35, MsWRKY51, MsWRKY52, MsWRKY67, MsWRKY85. Cluster 2 contained 10 MsWRKY transcripts, which were all significantly down-regulated at 1, 2 and 5 dpi, including MsWRKY7, MsWRKY12, MsWRKY32, MsWRKY70, MsWRKY71, MsWRKY94, MsWRKY101, MsWRKY102, MsWRKY103, MsWRKY105. In cluster 3, the expression levels of MsWRKY10, MsWRKY16, MsWRKY28, MsWRKY33, MsWRKY75, MsWRKY90 were significantly up-regulated at 1 dpi, and then down-regulated at 2 and 5 dpi, except for MsWRKY10 and MsWRKY33. MsWRKY13 and MsWRKY86 showed the same expression patterns, which were significantly increased at 2 and 5 dpi. The expression level of

![Figure 2. The conserved motif arrangements of MsWRKY proteins. (A) The phylogenetic tree of MsWRKY predicted proteins constructed by MAGA 7.0. (B) The motif compositions of WRKY proteins. The motifs 1–15 were displayed in different colors. (C) The sequence information of each motif.](image)
MsWRKY108 was significantly decreased at 1 and 2 dpi, and then were significantly increased at 5 dpi (Figure 4B). Overall, the results show that MsWRKYs might play important roles in the response to the *V. mali* infection in *M. sieversii*.

Figure 3. The specific transmembrane-binding domains of MsWRKY proteins. WRKY: WRYK domain; RPW8: RESISTANCE TO POWDERY MILDEW8.2 domain; rad50: rad50; UPF0242: UPF0242 domain; nuc_hydro: Nucleoside hydrolases; Plant_zn_clust: zinc-binding domain; PLN03210: PLN03210 domain.
The cut-off of the differentially expressed transcripts was $|\log_2 \text{(fold change)}| > 1$, $Q < 0.05$. Color scale represented the fold change of *V. mali* infection vs. 0 dpi control with normalized log2 (Fold change) values. Red, blue and white indicated high expression, low expression and no expression, respectively.

To confirm the contributions of the *MsWRKY* transcripts on *Valsa* canker resistance, twelve significantly differential expression *MsWRKY* transcripts in different groups were selected to quantify by qRT-PCR at different time points after infection. At 1 dpi, the relative expression level of *MsWRKY21, MsWRKY28, MsWRKY30* and *MsWRKY90* were significantly increased. At 5 dpi, the relative expression levels of *MsWRKY5, MsWRKY13, MsWRKY27, MsWRKY35, MsWRKY67, MsWRKY75, MsWRKY102* were significantly up-regulated. On the contrary, at 1 dpi the relative expression levels of *MsWRKY32* and *MsWRKY102* were significantly down-regulated. At 5 dpi, the relative expression level of *MsWRKY28* was significantly decreased. The qRT-PCR results of the twelve *MsWRKY* transcripts were consistent with the FPKM values from RNA-seq data (Figure 5).
3.6. Functional Identification of WRKY Family Members in M. sieversii Responded to V. mali

Dependent on the different expression patterns of these differentially expressed MsWRKY transcripts, six MsWRKY transcripts were cloned to identify the resistance to V. mali, including MsWRKY16, MsWRKY21, MsWRKY70, MsWRKY74, MsWRKY85 and MsWRKY105. To identify the potential resistance of MsWRKY proteins that responded to the infection of V. mali, we performed disease resistance verifications using the transient expression method mediated by A. tumefaciens [36]. Six MsWRKYs transcripts were separately transiently transfected into leaves using A. tumefaciens containing pBI121-MsWRKYs-GFP vectors. These relative expressions of MsWRKYs were correspondingly increased in leaves, after 4 days of the transiently overexpression by A. tumefaciens (Figure 6A), which indicated that these MsWRKY transcripts were successfully overexpressed in the leaves. In the dis-
ease resistance test, after 4-day inoculations of *V. mali*, the lesion ratios of the MsWRKY16, MsWRKY21, MsWRKY70 and MsWRKY85 were significantly decreased compared with the mock (Figure 6B). These results suggested that MsWRKY16, MsWRKY21, MsWRKY70, MsWRKY74 and MsWRKY85 could be positive regulators for the resistance to the *Valsa* canker. Yet, the mean lesion ratio of MsWRKY105 transiently expressed leaves was similar to the mock (Figure 6B), indicating that the MsWRKY105 was no resistant to the canker disease. Interestingly, MsWRKY70 and MsWRKY74 obviously induced severe necrosis at 6 days after transient expression compared with WT (Figure 6B).

Figure 6. The functional identification of WRKY family members in *M. sieversii* to response to the infection of *V. mali*. Six *MsWRKY* transcripts were transiently expressed in leaves of *M. domestica*. (A) The transfection efficiencies were determined by qRT-PCR at 4 days after transient expressions. The relative expressions were normalized with the reference gene *MsEF-1a*. (B) The *Valsa* canker disease resistance tests of transiently expressed *MsWRKYS* were presented. Mock: transiently expressed with *A. tumefaciens*. The lesion ratios were measured from detached leaves at 4 dpi. Each sample contained three biological replicates and each biological replicate contained three technical replicates. The means (bars: SE) were analyzed by ANOVA method, Tukey’s HSD test with significances comparing with the Mock (intra-group comparison). WT: the leaves were without the transient expressions; Mock: the leaves were transiently expressed with *A. tumefaciens* (without vectors). Lowercase letters presented significant differences (*p* < 0.05) in the Wound + *V. mali* group. The capital letter represented significant differences (*p* < 0.05) in the Wound + ddH2O group. The heatmap showing the qRT-PCR validation of *MsWRKY16*, *MsWRKY21*, *MsWRKY70*, *MsWRKY85* and *MsWRKY105* separately in *MsWRKYs* transiently expressed leaves in *M. domestica* under the phytohormone inductions and the *V. mali* infection at different time points (1 h, 3 h, 6 h, 1 d, 2 d and 5 d). The results are normalized to reference gene EF1-α and expressed as relative fold change compare to controls. The numbers in box represent the log2 of the relative expression with columns scale. The numbers in rackets represent standard errors (SEs). Phytohormones: SA (5 mM); MeJA (1 mM); ACC (1 mM). Each sample contained three biological replicates and each biological replicate contained three technical replicates. The means (bars: SE) were analyzed by ANOVA method, Tukey’s HSD test with significances comparing with the control (intra-group comparison). Lowercase letters presented significant differences (*p* < 0.05). The color bar indicates expression levels from low (blue) to high (red).
Plant hormones play important roles in signaling networks of immune defense involved in plant responses to pathogen infection [39]. To detect whether these six MsWRKY transcripts were induced by different phytohormones during the infection of *V. mali*, qRT-PCR was performed to determine the expression levels (Figure 6C). As the positive regulator for the canker disease resistance, the expression of MsWRKY16 was significantly up-regulated under the ACC + *V. mali*, MeJA + *V. mali* and the SA + *V. mali* conditions. With the ACC + *V. mali* treatment, the relative expression of MsWRKY16 significantly decreased at 6 h and 2 d. Under the MeJA + *V. mali* condition, the expression of the MsWRKY16 peaked at 6 h, yet was down-regulated after 2 d, as well as up-regulated at 5 d. The expression of MsWRKY16 was continuously increased after the SA + *V. mali* treatment. As the positive regulator of resistance to the infection of *V. mali*, MsWRKY21 was significantly up-regulated from 1 h to 5 d upon the treatment of MeJA + *V. mali*. Under the SA + *V. mali* and ACC + *V. mali* treatments, the MsWRKY21 was down-regulated from 1 h to 5 d compared to 0 h. The expression of MsWRKY70 was significantly down-regulated from 1 h to 2 d under the treatments of ACC + *V. mali*. and SA + *V. mali*. The expression of the MsWRKY85 was continuously down-regulated from 1 h to 5 d under the treatment of ACC + *V. mali*. Under the treatment of MeJA + *V. mali*, the expression of MsWRKY85 significantly increased from 1 h to 6 h, then decreased at 2 d. The expression of MsWRKY105 was continuously up-regulated from 1 h to 5 d in the treatment of MeJA + *V. mali*. Combining the results of the disease resistance trials, the MsWRKY16, MsWRKY21, MsWRKY70 and MsWRKY85 could play positive roles for the regulatory defense to the invasion of *V. mali* involving the ET, MeJA and SA signaling pathways.

4. Discussion

Wild apple is an ancestor of the cultivated apple distributed in Tianshan Mountain from Central Asia to West Europe along the Silk Road and holds underlying potential for the germplasm improvement of apple, dependent on genetic diversity [26,40]. However, *M. sieversii* was suffering from attacking by the canker disease caused by necrotrophic pathogen *V. mali*, resulting in a decrease in the area of this population [31]. The pathogen *V. mali* not only infects *M. sieversii*, but also causes extensive damage and even death to apples in eastern Asia with severe yield losses [41]. Thus, it is important to explore the disease resistance gene resources in *M. sieversii* upon the infection of *V. mali*, which could aid in the improvement of apple germplasm resources.

The WRKY family is among the ten largest families of transcription factors in higher plants and major positive or negative regulators in plant immunity [2]. WRKY families were identified in unicellular green alga *Chlamydomonas rein* (1), the moss *Physcomitrella patens* (37), *Arabidopsis* (74), *Glycine max* (197) and *Populus triuncatula* (103) [42]. In apple, based on the apple genome database of “Golden Delicious” [43], 116 WRKY transcription factor genes [44] and 119 MdWRKY genes in response to the fungal pathogen (*Alternaria alternata*) and phytohormone treatment [45] were separately identified. Recent research reported that a high-quality de novo assembly of the apple (GDDH13) was produced [4]. Based on this database, 114 WRKY transcription factor genes in *M. domestica* were identified [6]. In this study, a total of 112 MsWRKY transcripts were firstly identified from the full-length transcriptome data of *M. sieversii* upon the *V. mali* infection. The number of members of the WRKY family of *M. sieversii* identified from the full-length transcriptome responded to *V. mali* infection, is almost equal to the number in *M. domestica* (GDDH13) (Figure 1). Additionally, the number of members of the MsWRKY family expanded compared with *V. rein*, *P. patens*, *Arabidopsis* and *Populus*, suggesting that the number of WRKY genes has an expansion from lower plants to higher.

The phylogenetic analysis demonstrated that MsWRKYs could be classified into three groups, among which group II was further divided into five subgroups. The number of MsWRKYs group II accounted for the largest proportion of 67%, which was consistent with former reports regarding the WRKYs in *Arabidopsis*, *Populus* and *M. domestica* [2,4,6]. The NJ tree of MsWRKY proteins indicates that groups IIa and IIb, and groups IIId and Ile
have close phylogenetic relationships separately. The monophyletic clades of three groups seemed to be formed major five lineages regarding group I, group Ila + Iib, group IId +Ile, group Iic and group III (Figure 1). Fifteen conserved motifs were identified in the different groups of MsWRKYs (Figure 2B), indicating that the conserved motif composition of the group of WRKY family could be the cause of MsWRKY transcription factor with various biological functions. WRKY proteins are defined by the conserved WRKY domain, the amino acid sequence is WRKYGQK [7]. In a few WRKY proteins, the WRKY domain has some variants, including WRKYGKK, WRKYGEK, WRKYGMK and WKKYCEDK [46,47]. In this study, two variants were mainly observed in the subgroup Iic of MsWRKY proteins, which were WRKYGKK and WRKYGMK. It is inferred that MsWRKY50/MsWRKY61 and MsWRKY62, which possessed variants in WRKY domains (Table 1), might have a variety of biological functions. In N. tabacum, NtWRKY12 specifically binds the promoter of PR1a genes contained in theWK-box (TTTTCCAC), not W-box TTGAC(C/T), thereby improving disease resistance in the plant [48]. It could be inferred that MsWRKY50 and MsWRKY61 might have similar cis-acting elements, of which one is NtWRKY12. Furthermore, the biochemical properties of 112 MsWRKY proteins were characterized, including the grand average of hydropathicity (GRAVY), instability index, subcellular location, and phosphorylation site. The molecular weight of MsWRKY55 is the smallest of all, and consists of 80 amino acids. MsWRKY106 is the largest one of all, and consists of 995 amino acids. Twenty-seven MsWRKY proteins are without NLSs, may need other corresponding ligands to transport into the nucleus for biological functions (Table S1). MAP kinase pathways are involved in regulating WRKY transcription activation activity [16,49,50]. OsWRKY53 is a downstream target gene of OsMKK4-OsMPK3/OsMPK6, of which multiple clustered serine-proline residues (SP cluster) can be phosphorylated by OsMPK3/OsMPK6 to regulate the resistance to rice blast [51]. The abundant phosphorylation sites in MsWRKY proteins were observed (Table S1), speculating that they may play important roles in biological functions. The membrane-bound transcription factor is a kind of transcription factor with the C-terminal transmembrane domain (TMD). Under normal conditions, it localizes in the cytoplasm, while under stress conditions, the cleavage of the TMD allows the transcription factor to translocate to the nucleus [52,53]. In this study, four putative membrane-bound MsWRKY transcription factors were identified (Table S1). The conserve domain identification result showed that MsWRKY7, MsWRKY28, MsWRKY29, MsWRKY70–82 and MsWRKY106 had different specific conserved domains (Figure 3), inferring that these MsWRKY might not only be involved in transcriptional regulation as a transcription factor, but also play other roles in biological functions. Among these, the expression level of the MsWRKY7 transcript was significantly down-regulated at 5 dpi suggested that it might be involved in the response to V. mali infection (Figure 4B). The RPW8 domain can mediate defense resistance and cell death in the cell nucleus and cytoplasm, respectively, and finally drive broad-spectrum resistance to powdery mildew [54,55]. There are also two disease resistance RPW8 domains in the N-terminus of MsWRKY7 protein (Figure 3). It speculated that MsWRKY7 containing two RPW8 domains might also have the resistant function to powdery mildew.

The full-length RNA-seq transcriptome database was performed to profile MsWRKY transcripts expression patterns to respond to the Valsa canker disease. Thirty-four MsWRKY transcripts exhibited differential expression patterns upon V. mali infection (Figure 4B). Based on the expression patterns of the RNA-seq and qRT-PCR data, it is indicated that the differentially expressed MsWRKY transcripts may perform diversified functions in response to Valsa canker disease (Figure 5). Based on the different expression patterns, the selected six MsWRKY transcripts subsequently conducted the functional identification with a transient expressions assay. Among these, MsWRKY16, MsWRKY21, MsWRKY70, MsWRKY74 and MsWRKY85 were positive regulators in canker disease resistance (Figure 6). In the previous study, numerous WRKY genes have been identified as functions of disease resistances and clarified the mechanisms. AtWRKY18, AtWRKY40 and AtWRKY60 exhibit a response to the infection of P. syringae and B. cinerea [56]. In Arabidopsis, AtWRKY18 and AtWRKY40
play major roles in PAMP-triggered basal defense in defense to the powdery mildew (caused by fungus: *Golovinomyces orontii*) during the early stages of infection [57]. The elaborate functions and mechanisms of the canker disease resistance *MsWRKY* transcripts should be implemented for further research.

The WRKY family plays an important in the phytohormones synthesis and signaling pathway to immune defense. The co-regulation of AtWRKY33 and AtMYB46 mediates the coordination of SA and JA/ET pathways to orchestrate the defense against the phytopathogen [58]. The AtWRKY18, AtWRKY40 and AtWRKY33 act as repressors for PAMP-induced secreted peptide 3 (PIP3) to turn on the immunity via regulating SA and JA biosynthesis and signaling [59]. EF-Tu receptor (EFR) activates the SA and JA synthesis to defend against the infection, through the direct phosphorylation of a receptor-like cytoplasmic kinase (BIK1) and WRKY33, WRKY50 and WRKY57 [60]. Thus, the phytohormone-induced expression pattern of these disease resistance-related *MsWRKY* transcripts was verified. Upon the hormone treatments of ACC + *V. mali*, MeJA + *V. mali* and SA + *V. mali*, *MsWRKY16*, *MsWRKY21*, *MsWRKY70* and *MsWRKY85* showed different expression patterns (Figure 6C).

The expressions of *MsWRKY16* and *MsWRKY105* were induced through the treatments of MeJA + *V. mali* and SA + *V. mali*. The expressions of *MsWRKY21* and *MsWRKY85* were induced in the treatment of MeJA + *V. mali*. It speculated that these *MsWRKYs* could be key regulatory genes in the ET, JA and SA signaling pathways for defense of the *V. mali* infection. The potential functions and disease resistance regulatory mechanism of these *MsWRKY* TFs will be experimentally illuminated in our further research.

5. Conclusions

One hundred and twelve *MsWRKY* transcription factors were identified and divided into three groups (I, II, III) based on the phylogenetic analysis. Fifteen conserved motifs were identified in the different groups of *MsWRKYs*, of which compositions of similar motifs were usually clustered into the same clade. Based on the Illumina sequencing database, 34 differentially expressed *MsWRKY* transcripts may play key roles in response to the *V. mali* infection. Furthermore, the differentially expressed *MsWRKY16*, *MsWRKY21* *MsWRKY70*, *MsWRKY74* and *MsWRKY85* were identified to play positive roles in resistance to canker disease. Besides, these *MsWRKYs* might contribute to the regulation of resistance and susceptibility of *Valsa* canker disease involving the SA, MeJA and ET signaling pathways. These results will provide reliable indexes of screening candidate *MsWRKY* transcripts and build a foundation for further investigation in *M. sieversii* to respond to the *Valsa* canker disease.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/f12060790/s1, Table S1: The list of the 112 *MsWRKY* transcripts identified in this study with properties of the predicted *MsWRKY* proteins. Table S2: The expression patterns of differentially expressed *MsWRKY* transcripts. Table S3: List of the qRT-PCR primers of the differentially expressed *MsWRKY* transcripts.

Author Contributions: X.L. (Xiaojie Liu) and D.Z. conceived and planned the study. X.L. (Xiaojie Liu), Y.Z. and T.Z. completed the bioinformatics analysis and the qRT-PCR assay. X.L. (Xiaojie Liu), X.L. (Xiaoshuang Li), X.W. and D.Z. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (2019QZKK0502030403), the Science and Technology Plan Project of Xinjiang Uygur Autonomous Region (2016B03040) and the Youth Innovation Promotion Association, Chinese Academy of Sciences (No.2018478).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.
53. Sakuraba, Y.; Kim, D.; Han, S.H.; Kim, S.H.; Piao, W.; Yanagisawa, S.; An, G.; Paek, N.C. Multilayered regulation of membrane-bound ONAC054 is essential for abscisic acid-induced leaf senescence in Rice. *Plant Cell* **2020**, *32*, 630–649. [CrossRef]

54. Xiao, S.; Ellwood, S.; Calis, O.; Patrick, E.; Li, T.; Coleman, M.; Turner, J.G. Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. *Science* **2001**, *291*, 118–120. [CrossRef]

55. Huang, Y.Y.; Zhang, L.L.; Ma, X.F.; Zhao, Z.X.; Zhao, J.H.; Zhao, J.Q.; Fan, J.; Li, Y.; He, P.; Xiao, S.; et al. Multiple intramolecular trafficking signals in resistance to powdery mildew 8.2 are engaged in activation of cell death and defense. *Plant J. Cell Mol. Biol.* **2019**, *98*, 55–70. [CrossRef] [PubMed]

56. Xu, X.; Chen, C.; Fan, B.; Chen, Z. Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40, and WRKY60 transcription factors. *Plant Cell* **2006**, *18*, 1310–1326. [CrossRef]

57. Pandey, S.P.; Roccaro, M.; Schön, M.; Logemann, E.; Somssich, I.E. Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of *Arabidopsis*. *Plant J. Cell Mol. Biol.* **2010**, *64*, 912–923. [CrossRef]

58. Xie, M.; Zhang, J.; Yao, T.; Bryan, A.C.; Pu, Y.Q.; Labbe, J.; Pelletier, D.A.; Engle, N.; Morrell-Falvey, J.L.; Schmutz, J.; et al. *Arabidopsis* C-terminal binding protein ANGUSTIFOLIA modulates transcriptional co-regulation of MYB46 and WRKY33. *New Phytol.* **2020**, *228*, 1627–1639. [CrossRef] [PubMed]

59. Najafi, J.; Brembu, T.; Vie, A.K.; Viste, R.; Winge, P.; Somssich, I.E.; Bones, A.M. PAMP-induced secreted peptide 3 modulates immunity in *Arabidopsis*. *J. Exp. Bot.* **2020**, *71*, 850–864. [PubMed]

60. Lal, N.K.; Nagalakshmi, U.; Hurlburt, N.K.; Flores, R.; Bak, A.; Sone, P.; Ma, X.Y.; Song, G.Y.; Walley, J.; Shan, L.B.; et al. The receptor-like cytoplasmic kinase BIK1 localizes to the nucleus and regulates defense hormone expression during plant innate immunity. *Cell Host Microbe* **2018**, *23*, 485–497. [CrossRef] [PubMed]