J. Plant Biol. (2017) 60:215-230
DOI 10.1007/s12374-016-0577-3

ORIGINAL ARTICLE

Identification and Expression Analysis of *WRKY* Transcription Factor Genes in Response to Fungal Pathogen and Hormone Treatments in Apple *(Malus domestica)*

Shuai Lui[1], Changguo Luo[2], Longming Zhu[1], Renhe Sha[1], Shenchun Qu[1], Binhua Cai[1] and Sanhong Wang[1,*]

[1]*College of horticulture, Nanjing Agricultural University 210095, Nanjing, China*
[2]*Guizhou fruit institute, Guizhou Academy of Agricultural Science 550006, Guiyang, China*

Received: November 25, 2016 / Accepted: December 29, 2016

Abstract The WRKY family, a large family of transcription factors (TFs) found in higher plants, plays central roles in many aspects of biological processes and adaption to environment. However, little information is available on this family in apple (*Malus domestica*). In the present study, a total of 119 candidate *WRKY* genes in apple genome were identified and classified into three main groups (Group I–III) based on the structure of the conserved domains. Each group or subgroup showed similar exon–intron structures and motif compositions. The evolution analysis showed that 44 *MdWRKY* genes clustered into 20 intensive regions (<100 kb) and 78 *MdWRKY* formed 85 pairs of collinear relationships, suggesting that both tandem and segmental duplications played an important role in the evolution and diversification of the WRKY gene family in apple. Furthermore, the expression of the *MdWRKY* genes in apple leaves in response to biotic stress (*Alternaria alternate*) and hormone treatments [salicylic acid (SA), methyl jasmonate (MeJA) and ethephon] was examined by using RNA-seq and qRT-PCR. The results showed that 63 *MdWRKY* genes had differential expression in their transcript abundance in response to *Alternaria alternata* apple pathotype infection. Moreover, most pathogen responsive *MdWRKY* genes were also changed significantly when apple leaves were treated by SA, MeJA or ethephon plant growth regulations, suggesting an interaction between SA, JA and ethylene (Eth) hormone signaling under biotic stress. This work may provide the basis for future studies of the genetic modification of WRKY genes for pathogen resistance in apple.

Keywords: *Alternaria alternata* apple pathotype, Apple, Bioinformatics analysis, Expression profile analysis, Hormone treatment, WRKY transcription factor

*Corresponding author; Sanhong Wang
Tel : +86-25-84395724
E-mail : wsh3xg@hotmail.com

Introduction

Plants are affected by a wide range of environmental stresses during their lifetime due to their inability to escape harmful environmental conditions. Plants have evolved subtle strategies to adapt to and deal with biotic and abiotic stresses, which are mediated by a complex signal transduction network that results in changes in biochemistry, physiology and morphology. Transcription factors (TFs) play important roles in regulation and control many crucial biological processes by regulating gene expression. WRKY TFs are a large family of regulatory proteins. They exist primarily in plants, but absent in animals with the exception of several examples in protozoa (Zhang and Wang 2005). Since the first cDNA coding a WRKY protein (SPF1) was cloned from sweet potato (Ishiguro and Nakamura 1994), a large number of WRKY proteins have been identified from various classes of plants, such as Arabidopsis (Eulgem et al. 2000), rice (Wu et al. 2005), barley (Mangelsen et al. 2008), coffee (Ramiro et al. 2010), cucumber (Ling et al. 2011), grape (Wang et al. 2014), cotton (Cai et al. 2014), *Populus* (Jiang et al. 2014).

WRKY TF contains one or two highly conserved WRKY domain(s) of approximate 60 amino acids with the conserved sequence WRKYGQK followed by a zinc-finger-like motif C_2H_2 ($C–X_4–C–X_{22–23}–H–X_1–H$) or C_2HC ($C–X_7–C–X_{23}–H–X_1–C$) (Eulgem et al. 2000). This domain generally binds to the DNA element termed W-box ((C/T)TGAC(T/C)) of their target genes (Eulgem et al. 2000; Ciolkowski et al. 2008), although alternative binding sites such as SURE or SURE-like element have been also identified (Sun et al. 2003; Rushton et al. 2010). Based on the number of WRKY domains and certain features of the zinc-finger motif, WRKY proteins

can be classified into three groups (Group I-III). Proteins contain two WRKY domains (in N-terminal and C-terminal, respectively) belong to Group I, whereas proteins with one WRKY domain belong to Group II or Group III. Generally, the WRKY domains in Group I and Group II have the zinc finger motif, C_2H_2, while the single WRKY domain in Group III is different from that in Group I and Group II, the WRKY domain containing a C_2HC motif, rather than a C_2H_2 pattern. Moreover, Group II proteins can be further subdivided into five subgroups (IIa, IIb, IIc, IId and IIe) based on additional short conserved structural motifs (Eulgem et al. 2000; Zhang and Wang 2005; Fan et al. 2011).

Plant WRKY TFs not only play pivotal roles in regulating many aboitic stress reactions, such as salinity (Qin et al. 2013), drought (Tripathi et al. 2014), heat (Li et al. 2011) and cold (Zhou et al. 2008; Zou et al. 2010), but also in regulating defense-related genes and enhancing disease resistance to pathogens (Rushton et al. 2010). In a number of plants, most WRKY genes are rapidly induced by pathogens, pathogen elicitors, or SA treatment (Eulgem et al. 1999; Chen and Chen 2000; Dong et al. 2003; Ryu et al. 2006; Ramamoorthy et al. 2008). Over-expression of *AtWRKY6*, *AtWRKY18*, *AtWRKY53* and *AtWRKY70* results in enhanced disease resistance to pathogens and induction of defense-related genes (Robatzek and Somssich 2001; Chen 2002; Ulker et al. 2007; Murray et al. 2007). In contrast, constitutive expression of *AtWRKY11*, *AtWRKY17*, *AtWRKY25* and *AtWRKY33* support enhanced growth of pathogen, suggesting these WRKY proteins are negative regulators in defense signaling (Journot-Catalino et al. 2006; Zheng et al. 2007; Zheng et al. 2006). Accumulated evidences showed that WRKY TFs regulated plant defense response by phytohormone (SA, JA and Eth) signaling. For instance, overexpression of *AtWRKY75* enhanced Arabidopsis resistance to *Pectobacterium carotovora* ssp. *Carotovora* (*Pcc*) mediated by SA and JA defense signaling (Choi et al. 2014). Overexpression of *AtWRKY28* and *AtWRKY75* enhanced resistance to *Sclerotinia sclerotiorum* and oxalic acid stress in Arabidopsis mainly by JA/Eth pathway (Chen et al. 2013). *AtWRKY33* plays opposite roles in regulation of Arabidopsis defense response to necrotrophic pathogens and biotrophic pathogen mediated by JA and SA defense signaling pathways, respectively (Zheng et al. 2006). *Capsicum annuum* WRKY proteins CaWRKY27, CaWRKY40 act as positive regulators in tobacco resistance responses to *R. solanacearum* infection through SA, JA and Eth mediated signaling pathways (Dang et al. 2013; Dang et al. 2014). *PtrWRKY89* is a regulator of a SA-dependent defense signaling pathway in poplar (Jiang et al. 2014).

Apple (*Malus* × *domestica*) is cultivated widely and has a great economic importance. The apple productivity is severely affected by various environmental stresses. Some research showed that WRKY TFs play important roles in apple response to abiotic and biotic stress and regulation of resistance. For example, overexpression of *MdWRKY13* from apple (*Malus* × *domestica*) in Arabidopsis resulted in drought tolerance decrease (Duan et al. 2014). *MdWRKY1* was induced dramatically by *Alternaria alternata* apple pathotype (formerly named *Alternaria alternata f.* sp. mali) infection and also by SA and MeJA treatments. Overexpression of *MdWRKY1* in tobacco enhanced resistance to *Phytophthora parasitica* var. *nicotianae* Tucker (Fan et al. 2011). Recently, Meng et al (2016) identified WRKY genes in apple by whole genome analysis and analyzed their responses to waterlogging and drought stress. However, there are few reports regarding on apple WRKY TFs in response to pathogen infection. Alternaria leaf blotch, one of the most damaging leaf spot diseases of cultivated apple (Bulajic et al. 1996), is caused by the fungal pathogen *A. alternata* apple pathotype (AP). Based on the important roles of WRKY TFs in the regulation of transcriptional reprogramming associated with plant immune responses, we would like to know the *A. alternata* induced WRKY genes in apple and their potential functions in response to the pathogen infection. In this study, a total of 119 *WRKY* genes were identified in apple and the gene classification and phylogenetic relationship were performed by whole-genome analysis. The expression profiles of the *MdWRKYs* in response to *A. alternata* AP infection and under hormone treatments (SA, MeJA and ethephon) were analyzed by RNA-seq and quantitative real-time PCR (qRT-PCR). These results provided a foundation for further studying the function of MdWRKY TFs in apple disease resistance response.

Results

Genome-wide Exploration of *MdWRKY* Genes and Phylogenetic Analysis

The prominent structural feature of WRKY proteins is WRKY domain, which consists of a highly conserved sequence of WRKYGQK, followed by a zinc-finger motif (Eulgem et al. 2000). For downloading sequences of *WRKY* genes in the apple genome, we used a HMMER-BLASTP-InterProScan strategy to search for genes encoding proteins containing the Pfam PF03106 domain. Based on manual inspection using the MEGA 6.0 software, we discarded the sequences which contain insufficient conservation of WRKY sequences. A total of 119 predicted *MdWRKY* genes were identified and named *MdWRKY1a* ~ *MdWRKY75f* based on the homology with *AtWRKYs*. Among the 119 *MdWRKY* genes, 115 are distributed across the 17 apple chromosomes except 4 unanchored to any chromosomes (Table 1, Fig. 1). Chromosomes 1, 9, and 12 contain relatively more *MdWRKY* genes, with 11, 11 and 12 genes, respectively, whereas chromosomes 2 and 5 contain the

Table 1. *WRKY* genes and related information in apple

Gene name	Gene ID	Chromosome	Start	End	Strand	Len	MW	pI	E-value
MDP0000256105	MdWRKY1a	chr9	8888150	8890196	+	484	53.22	6.37	3.80E-23
MDP0000260803	MdWRKY1b	chr17	9013295	9015664	+	485	53.33	7.14	6.70E-26
MDP0000648338	MdWRKY2a	chr3	4161709	4164777	+	733	79.64	6.39	9.90E-25
MDP0000144203	MdWRKY2b	chr4	22672804	22676168	+	726	79.72	6.51	6.00E-21
MDP0000184044	MdWRKY2c	chr12	31340923	31345960	+	965	106.62	7.4	1.90E-22
MDP0000258212	MdWRKY3a	chr16	2725244	2727634	+	562	61.16	9.52	5.00E-26
MDP0000849514	MdWRKY3b	chr13	4205206	4207916	-	526	57.34	7.71	4.20E-26
MDP0000179145	MdWRKY3c	chr9	3568378	3571919	+	528	57.63	8.42	2.70E-25
MDP0000242596	MdWRKY3d	chr9	3537891	3546179	+	1233	136.24	9.52	7.90E-25
MDP0000598428	MdWRKY3e	chr12	17836814	17837469	+	80	9.32	10.51	3.40E-25
MDP0000125782	MdWRKY4a	chr13	4152851	4155122	+	436	47.33	6.44	9.40E-07
MDP0000128464	MdWRKY4b	chr17	3858452	3860259	-	342	36.74	9.75	4.80E-25
MDP0000935652	MdWRKY6a	chr10	32001774	32004275	+	611	65.63	6.95	4.80E-26
MDP0000301666	MdWRKY6b	chr5	1123886	1126451	-	629	68.06	7.62	5.00E-26
MDP0000129882	MdWRKY7a	chr8	12838197	12839860	+	351	38.12	9.97	1.80E-26
MDP0000231668	MdWRKY7b	chr15	2170222	2171733	+	342	37.17	9.52	8.60E-27
MDP0000130716	MdWRKY7c	unanchored	100473904	100475135	-	330	35.99	10.15	8.00E-27
MDP0000128463	MdWRKY7d	chr2	16825270	16826501	-	330	35.97	10.15	9.40E-27
MDP0000179719	MdWRKY9a	chr8	28098537	28100933	+	572	62.54	5.23	2.30E-25
MDP0000137704	MdWRKY9b	chr15	43865669	43869894	-	606	67.87	5.72	1.90E-26
MDP0000272940	MdWRKY11a	chr15	4306158	4307696	-	338	36.71	9.94	6.00E-27
MDP0000185288	MdWRKY11b	chr8	15912158	15913661	+	341	37.06	9.67	7.30E-27
MDP0000716551	MdWRKY11c	chr8	15910975	15913994	-	382	41.28	9.76	7.40E-27
MDP0000130400	MdWRKY11d	unanchored	3297390	3298491	+	280	30.53	10.47	6.10E-26
MDP0000946614	MdWRKY11e	chr3	22215461	22216527	-	281	30.72	10.57	1.90E-26
MDP0000381897	MdWRKY11f	chr15	4126920	4128285	+	101	11.68	10.09	3.60E-25
MDP0000259279	MdWRKY13a	chr8	25743083	25746605	+	270	30.43	8.95	2.40E-25
MDP0000176224	MdWRKY13b	chr14	6054931	6059446	+	320	35.41	9.51	2.70E-12
MDP0000566005	MdWRKY14a	chr10	24876832	24887045	-	537	58.22	6.88	8.80E-26
MDP0000168871	MdWRKY14b	chr14	10894491	10899025	-	482	52.61	6.17	6.80E-26
MDP0000200748	MdWRKY17	chr15	4123429	4125687	-	164	18.31	9.55	1.20E-24
MDP0000184361	MdWRKY20a	chr17	9258377	9263739	-	692	75.91	7.13	2.70E-24
MDP0000289397	MdWRKY20b	chr11	21845938	21850053	-	616	67.15	6.4	6.80E-23
MDP0000294643	MdWRKY20c	unanchored	98677186	98679778	-	342	37.58	7.52	1.00E-24
MDP0000281965	MdWRKY21a	chr10	19648159	19651418	-	437	49.09	9.87	1.40E-22
MDP0000322257	MdWRKY21b	chr6	11443530	11449618	+	963	108.09	7.32	4.80E-17
MDP0000909869	MdWRKY22a	chr7	13388092	13389364	+	348	37.69	7.24	2.90E-26
MDP0000602139	MdWRKY22b	chr1	9738347	9739627	-	339	36.48	7	3.00E-27
MDP0000202292	MdWRKY22c	chr11	34912048	34913409	-	268	30.82	4.54	1.60E-25
MDP0000294489	MdWRKY22d	chr3	32982869	32984216	+	260	30.33	4.87	3.00E-24
MDP0000134105	MdWRKY23a	chr9	31171605	31173048	+	350	38.41	6.23	2.50E-27
MDP0000127976	MdWRKY23b	chr9	31182729	31184171	-	350	38.39	6.23	2.50E-27
MDP0000657441	MdWRKY23c	chr17	24207876	24219074	+	384	42.15	6.94	5.80E-27
MDP0000652760	MdWRKY23d	chr9	31177401	31178127	+	186	20.54	8.81	6.90E-28
MDP0000294526	MdWRKY27a	chr1	27340620	27342989	+	606	66.71	6.42	6.90E-26
MDP0000135668	MdWRKY27b	chr7	24496187	24497746	+	433	48.55	6.51	5.00E-26
MDP0000119590	MdWRKY28a	chr10	26140734	26142417	-	327	35.89	6.95	2.90E-26
MDP0000118810	MdWRKY28b	chr5	6341860	6343547	+	319	35.11	7.11	3.00E-26
MDP0000257129	MdWRKY28c	chr17	11890140	11892494	-	365	41.25	7.27	2.70E-26
MDP0000505247	MdWRKY28d	chr9	11591755	11599818	-	452	50.52	8.25	3.40E-24
MDP0000247896	MdWRKY29a	chr7	14142892	14144253	+	316	35.94	4.94	3.10E-25
MDP0000145953	MdWRKY29b	chr2	33651152	33652415	-	254	28.81	7.37	1.10E-25
MDP0000131218	MdWRKY32a	chr15	8129660	8132989	-	483	53.34	6.86	9.10E-24

Table 1. Continued

Gene name	Gene ID	Chromosome	Start	End	Strand	Len	MW	pI	E-value
MDP0000293456	MdWRKY32b	chr2	1463114	1466139	+	477	52.55	8.72	1.20E-18
MDP0000195385	MdWRKY33a	chr12	25273072	25275034	-	509	56.48	7.22	4.60E-26
MDP0000296025	MdWRKY33b	chr12	25272427	25274387	-	512	56.82	7.22	1.30E-24
MDP0000514115	MdWRKY33c	chr3	5212258	5214802	+	571	62.55	7.5	8.70E-26
MDP0000708692	MdWRKY33d	chr4	16750826	16752798	-	520	57.71	7.6	1.20E-25
MDP0000935996	MdWRKY33e	chr11	5219008	5221576	+	572	62.9	7.28	1.10E-25
MDP0000507805	MdWRKY33f	chr3	5214648	5216866	-	464	51.63	8.7	6.90E-26
MDP0000177906	MdWRKY40a	chr17	17980731	17982494	+	321	35.67	8.29	6.40E-24
MDP0000794439	MdWRKY40b	chr14	14574950	14576650	+	320	35.31	7.83	1.60E-24
MDP0000689162	MdWRKY40c	chr1	26539183	26540631	+	303	33.73	7.5	1.30E-23
MDP0000307516	MdWRKY40d	chr8	7696856	7698700	+	334	37	7.43	1.10E-23
MDP0000263349	MdWRKY40e	chr1	26554166	26555673	+	286	32.09	8.11	3.00E-24
MDP0000228838	MdWRKY40f	chr8	7673200	7674531	+	278	31.26	8.8	1.40E-23
MDP0000263961	MdWRKY40g	chr8	7681736	7686175	+	352	39.5	8.83	1.80E-23
MDP0000299114	MdWRKY41a	chr1	23915631	23917202	-	351	39.44	6.01	2.60E-26
MDP0000146360	MdWRKY41b	chr1	23902386	23903957	-	351	39.44	6.01	2.60E-26
MDP0000315045	MdWRKY41c	unanchored	9473875	9475683	-	355	39.87	5.86	1.30E-25
MDP0000767097	MdWRKY41d	chr7	24835032	24836577	-	342	38.05	5.59	3.20E-26
MDP0000288378	MdWRKY43	chr1	9685197	9691219	-	339	38.47	10	4.50E-20
MDP0000169621	MdWRKY44a	chr4	10893937	10896214	+	471	51.68	9.1	5.30E-26
MDP0000268364	MdWRKY44b	chr12	17810585	17813522	-	470	51.55	9.15	5.50E-26
MDP0000201945	MdWRKY44c	chr12	17825194	17826756	-	240	26.28	6.51	4.60E-26
MDP0000788581	MdWRKY44d	chr4	10921722	10922563	+	146	16.66	10.06	1.20E-25
MDP0000263900	MdWRKY45	chr14	21681842	21682591	+	135	15.58	10.09	1.30E-25
MDP0000455180	MdWRKY46a	chr1	10820928	10822933	+	347	38.46	5.05	8.00E-25
MDP0000727570	MdWRKY46b	chr1	10818735	10820741	-	347	38.46	5.05	8.00E-25
MDP0000146390	MdWRKY48a	chr13	12501144	12502874	+	385	42.67	6.39	1.40E-27
MDP0000256514	MdWRKY48b	chr16	8325971	8327657	-	371	41.14	5.8	3.30E-27
MDP0000205962	MdWRKY49a	chr12	19511944	19513835	-	297	33.11	6.26	9.70E-26
MDP0000215371	MdWRKY49b	chr4	12772602	12774493	+	315	35.3	5.99	1.20E-18
MDP0000255768	MdWRKY49c	chr12	19511741	19519340	+	951	105.34	8.7	4.90E-25
MDP0000121669	MdWRKY50a	chr6	5765161	5766017	+	161	18.24	5.72	4.80E-25
MDP0000299555	MdWRKY50b	chr8	8485908	8487854	+	252	28.58	9.25	1.60E-18
MDP0000253189	MdWRKY51	chr17	18303282	18306886	+	217	24.3	9.42	6.10E-24
MDP0000219647	MdWRKY53a	chr1	27704346	27705878	-	347	38.78	6.43	5.40E-25
MDP0000468391	MdWRKY53b	chr1	27704241	27705773	-	347	38.76	6.32	5.40E-25
MDP0000191017	MdWRKY53c	chr7	16473459	16475677	+	356	39.73	5.57	8.20E-25
MDP0000460139	MdWRKY55a	chr12	26275509	26278113	+	344	37.66	6.86	1.40E-26
MDP0000828055	MdWRKY55b	chr4	3332748	3334900	-	342	37.44	5.58	1.10E-26
MDP0000228328	MdWRKY56a	chr7	13320138	13321638	+	222	24.93	9.73	3.70E-26
MDP0000119031	MdWRKY56b	chr7	13350689	13352189	-	222	24.93	9.73	3.70E-26
MDP0000707539	MdWRKY57a	chr16	2607579	2609994	-	246	26.65	8.48	1.40E-25
MDP0000193216	MdWRKY57b	chr13	3949136	3952405	+	398	43.79	7.9	4.10E-25
MDP0000418900	MdWRKY60	chr9	9238758	9239305	-	158	17.43	8.74	4.60E-24
MDP0000196330	MdWRKY61a	chr11	10697561	10700834	+	585	63.33	6.79	1.80E-25
MDP0000416279	MdWRKY61b	chr11	10703622	10706895	+	585	63.44	6.79	1.80E-25
MDP0000133918	MdWRKY65a	chr10	27085544	27086768	+	266	29.52	4.89	3.10E-26
MDP0000304113	MdWRKY65b	chr5	5629399	5630605	+	273	30.24	6.18	1.50E-25
MDP0000676216	MdWRKY69	chr9	26720765	26727339	-	354	38.46	8.56	3.30E-25
MDP0000175240	MdWRKY70a	chr12	26279186	26282001	-	303	34.31	6.05	4.00E-23
MDP0000228304	MdWRKY70b	chr4	3328450	3331822	+	362	40.82	6.84	6.60E-12
MDP0000754989	MdWRKY70c	chr12	26320212	26328572	+	455	51.62	7.04	1.50E-22
MDP0000290374	MdWRKY72a	chr17	3237713	3239968	+	455	49.95	7.75	4.60E-25

Table 1. Continued

Gene name	Gene ID	Chromosome	Start	End	Strand	Len	MW	pI	E-value
MDP0000452559	MdWRKY72b	chr9	2872733	2874447	-	412	44.78	7.98	8.70E-25
MDP0000308261	MdWRKY72c	chr6	20805475	20810729	-	707	76.67	6.82	7.10E-26
MDP0000300712	MdWRKY72d	chr6	20752772	20758054	-	738	80.46	7.14	7.50E-26
MDP0000161881	MdWRKY72e	chr16	3381593	3384229	-	593	64.87	9.01	4.30E-26
MDP0000273851	MdWRKY72f	chr13	4957184	4960950	+	576	62.35	6.47	9.60E-20
MDP0000231993	MdWRKY74a	chr4	21489801	21491070	+	325	36.52	10.3	4.50E-26
MDP0000234335	MdWRKY74b	chr12	30245947	30247495	-	319	35.86	10.79	1.40E-24
MDP0000123888	MdWRKY75a	chr16	6561225	6563173	-	190	21.4	10.34	2.70E-26
MDP0000142583	MdWRKY75b	chr17	612710	614588	-	223	25.6	9.43	2.50E-26
MDP0000792088	MdWRKY75c	chr13	8495422	8497092	+	226	25.55	10.15	1.20E-07
MDP0000154734	MdWRKY75d	chr9	677443	680156	-	233	26.71	9.51	4.30E-07
MDP0000263768	MdWRKY75e	chr6	17203132	17203909	-	150	17.15	10.01	2.00E-25

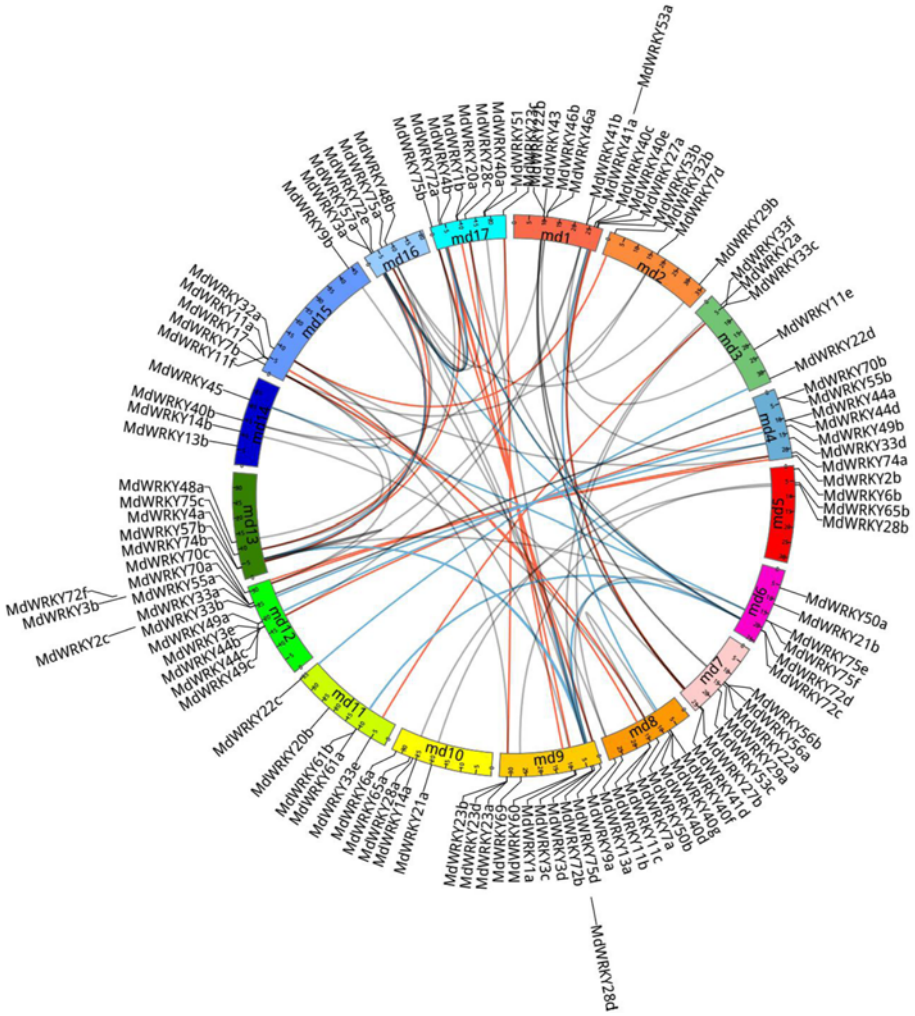

Fig. 1. Localization and duplication of the *WRKY* genes in the apple genome. Circular visualization of the *WRKY* genes was mapped on the different chromosomes in the genome using the Circos software. Chromosome number is indicated on the chromosome. The synteny relationship between each pair of *WRKY* genes were detected by using the MicroSyn software. The genes have synteny relationship are linked by lines. Red link: >30 anchors in a synteny block, blue link: 20-30anchors, green link: 10-20 anchors, gray link: 5-10 anchors.

fewest with only 3 genes each. Putative proteins encoded by *MdWRKY* genes contain 80 ~ 1233 amino acids, with the pI ranged from 4.54 (*MdWRKY22c*) to 10.79 (*MdWRKY74b*). The information of all *MdWRKY* genes was listed in Table 1. The range of variability implies that different MdWRKY TFs might operate in different microenvironments.

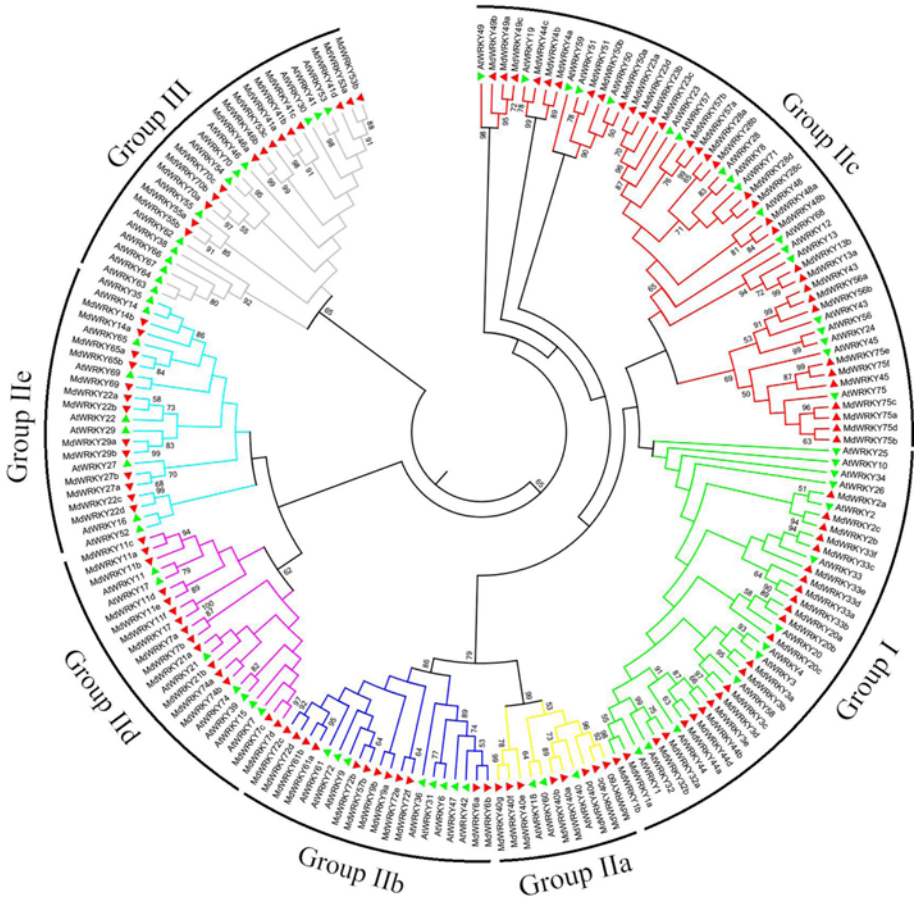

Fig. 2. Phylogenetic trees of WRKY proteins in apple and Arabidopsis. The un-rooted phylogenetic tree of *WRKY* full length protein sequences was constructed with MEGA6 program with the neighbor-joining method. The numbers beside the branches represent bootstrap values based on 1000 replications. Branch lines of subtrees are coloured indicating different WRKY subgroups.

Since the WRKY family in Arabidopsis was well studied and many members have been extensively investigated (Eulgem et al. 2000), we constructed a phylogenetic tree based on the alignment of WRKY domains from apple and Arabidopsis to evaluate the phylogenetic relationships and classified them within the established subfamilies (Fig. 2). All *MdWRKY* genes were classified into three groups. Twenty-four members were classified to Group I, which containing two WRKY domains and the C_2H_2-type zinc-finger motifs. Eighty-one proteins possessing a single WRKY domain were assigned to Group II in which also contains C_2H_2-type zinc-finger structure. The Group II was further divided into five subgroups IIa (8), IIb (12), IIc (33), IId (15) and IIe (13) as described by Eulgem et al. (2000). Three proteins (MdWRKY50a, -51 and -50b) in Group IIc were found containing WRKYGKK domain, rather than WRKYGQK which is universal in other WRKY proteins (Fig. S1). Fourteen proteins that have one WRKY domain but C_2HC zinc-finger motif were classified to Group III.

Exon–intron Structure and Motifs of the MdWRKY Gene Family

To gain more insight into the evolution of the WRKY family in apple, we analyzed the exon-intron structure of *MdWRKY* genes. As shown in Fig. 3, each groups mostly shared the similar exon-intron structural pattern. One hundred and two *MdWRKY* genes contained one to four introns and 17 had five to twelve introns. *MdWRKY3d* (in Group I) contained 12 intron, the largest number introns than other genes contained. The WRKY domain (in Group II, III) and C-terminal WRKY domain (in Group I) of MdWRKY proteins contained one intron except for MdWRKY29b, MdWRKY29a and MdWRKY4b proteins containing no introns (Fig. 3). In addition, other motifs also exist in MdWRKY proteins besides WRKY domains. The top 10 motifs in MdWRKY proteins were surveyed by using MEME program (Bailey et al. 2009), which ranged from 21 to 41 residues and distributed in each protein differently (Fig. 4, Table 2). Moreover, similar motif compositions were found within the same subgroups. Motifs 1, 2 and 5 represented the conserved motifs of the WRKY domain. Motif 9 is unique to Group III, whereas motif 10 only exists in Group IIa and IIb. One or more conservative motifs outside of the WRKY domain were detected manually in the

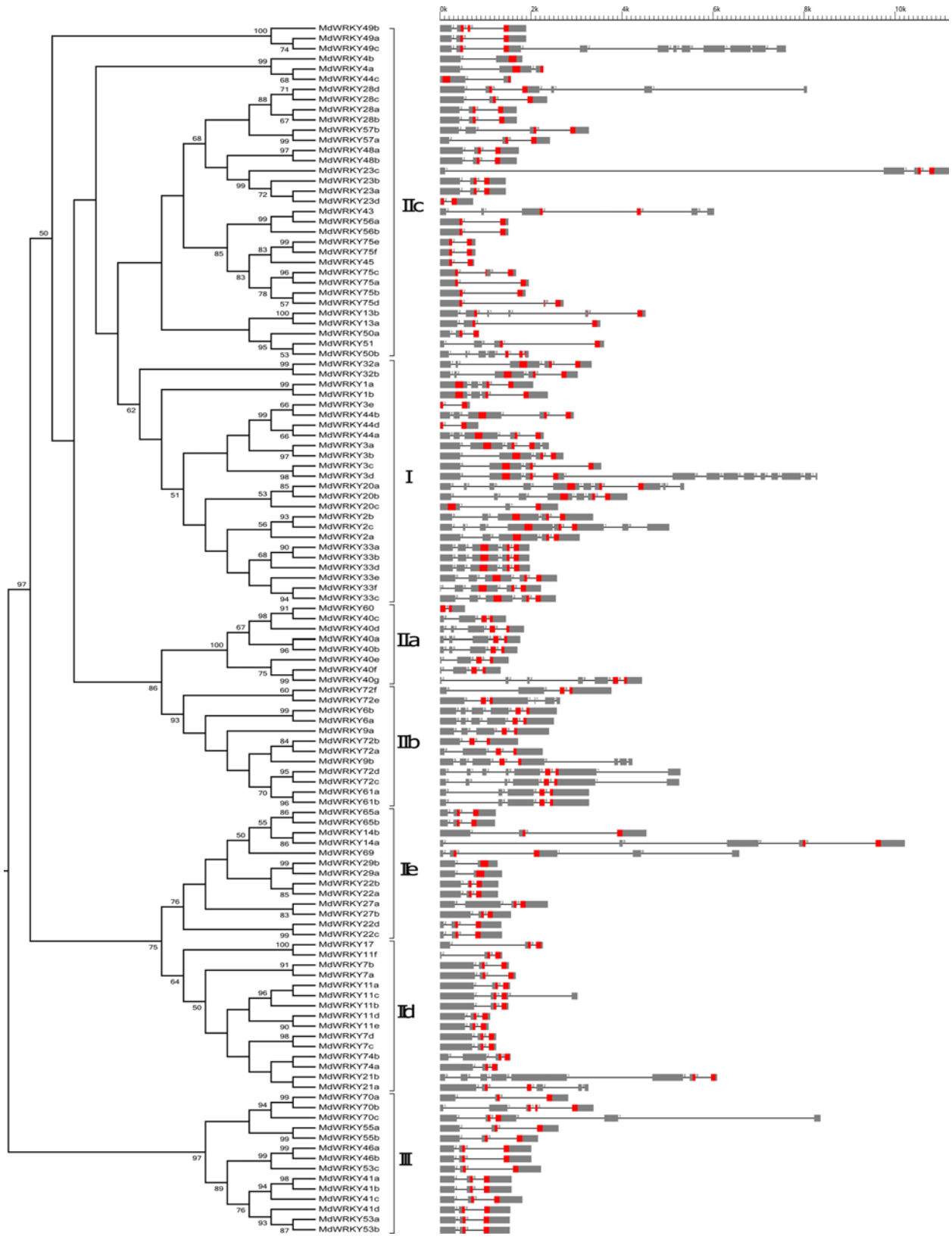

Fig. 3. Schematic representations of the exon–intron compositions of *MdWRKY* proteins from different groups or subgroups. Phylogenetic tree of *WRKY* genes are placed on the left. Exons, represented by boxes, were drawn to scale. Dashed lines connecting two exons represent an intron. Intron phases 0, 1 and 2 are indicated by numbers 0, 1 and 2, respectively. WRKY domain in MdWRKY proteins is marked in red.

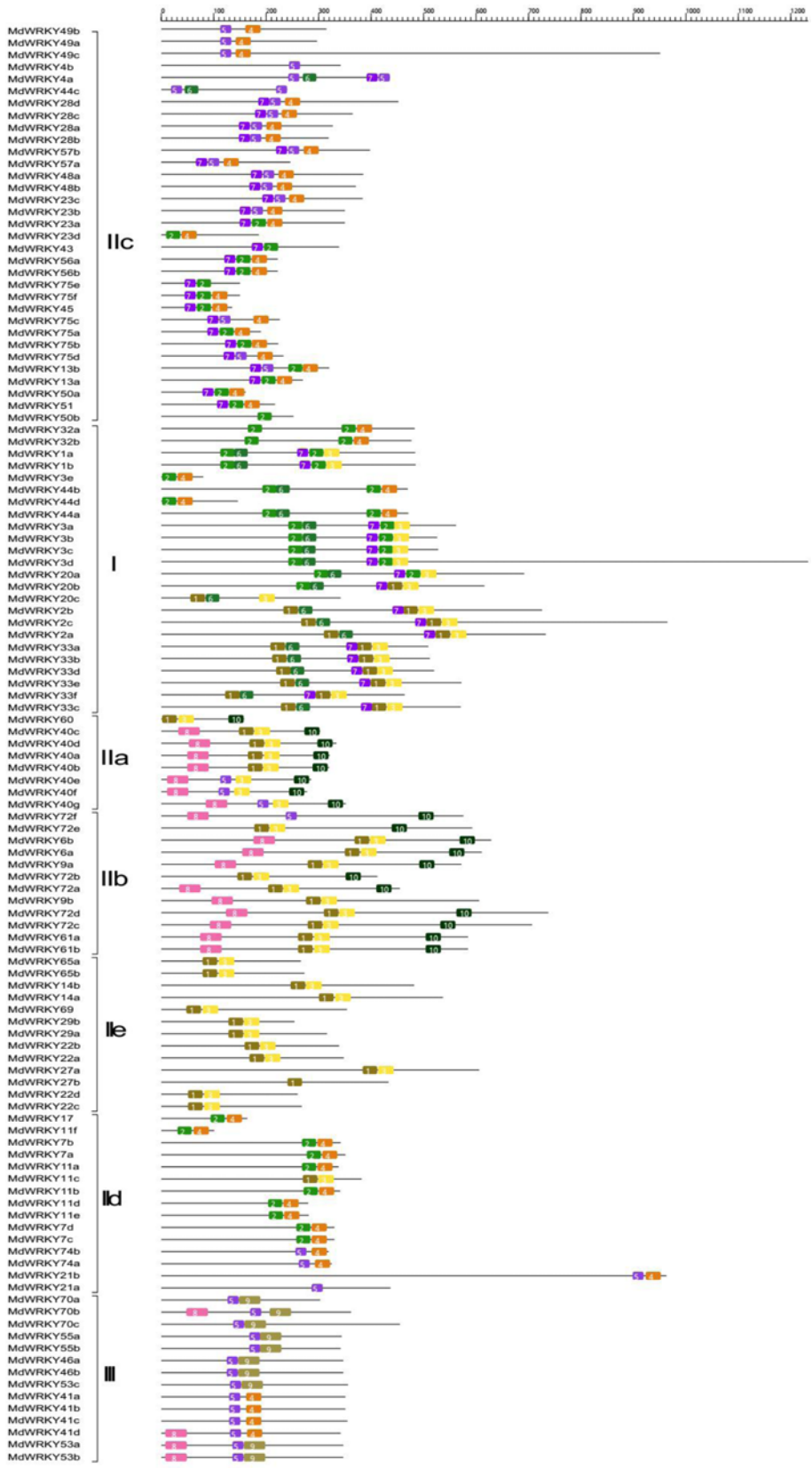

Fig. 4. Schematic representations of the conserved motifs of *MdWRKY* proteins from different groups or subgroups. Names of genes are indicated on the left. Different motifs are highlighted with different colored boxes with numbers 1 to 10. Lines represent protein regions without detected motif.

Table 2. Conserved motifs in MdWRKYs

Motif Number	Width	Best possible match
1	28	[IT][SL][DNS]DG[YWC][QANR]WRKYGQK[VP][VI]KG[NS][PE][NCY]PR[SAG]YY[KR]C
2	27	DDGY[RN]WRKYGQK[VP][VI]K[GN][SN]P[YFH]PR[SG]YY[KR]C[TS]
3	30	[SA][PKA][GS]CP[VA][RK]K[QH]V[EQ]R[ASC]S[EH]DP[SK][IA][LV][IV]TTYEG[EK]HNH
4	29	[GK]C[NP][VA][KR]K[QRH]V[EQ]R[SAL]S[DE]DPS[MIT][VL][IV][TV]TYEG[QE]H[NT]H[PS]
5	21	[DE]DGY[RS]WRKYGQKA[IV][KL][GN][SA][PK][YFH]PR
6	26	HPNC[PE][VT]KKKVERSLDGQITEI[IV]YKG[EN]
7	21	KKGx[KR]KV[RK]EPR[FV][AV][FV]QT[KRT]SEVD
8	41	[EQ]LDS[LA]Q[EA]E[LM][GQ][RE]VREEN[QK]RL[KR][EK][ML]Lx[VQ][IM]T[EK][DN]Y[QNS]TL[QE]MQ[LFV]xDIM
9	41	[GA]Y[YF]RCT[HR][RKQ][GKNY][DTLV][QY][GNS]C[KV]A[TK]K[QH]VQ[RKQ][SIL][DQ][DAE][DN]P[TLS][TI][FY][EVQ][ILTV]TY[RI]G[EKV]HTC[RHNT][QKM]
10	29	[VI][AE][DQA][AM][TVA][KS][AS][IL]T[SK]DP[NSD]F[TQ]SALAAA[IL][TS][SG][IR]IG[GN][GHN]

Motif numbers correspond to the motifs in Fig. 4

MdWRKY proteins. HARF (RTGHARFRR[A/G]P) (Eulgem et al. 2000), a conserved calmodulin-binding motif (Fan et al. 2011), was found in nine members of Group IId. Thirty-three proteins contain LxxLL motif (or LxLxLx motif) (Tiwari et al. 2004), which has previously been proven to be important to immune responses in plants (Rairdan and Moffett 2007). Eight of the thirty-three proteins (MdWRKY72e, -49c, -40e, -40f, -40g, -43, -21b and -75c) contained both LxxLL and LxLxLx motifs (Supplementary Table S3). The variety and complexity of these motifs in MdWRKY proteins indicate that MdWRKY TFs may play important roles in the growth regulation and stress responses.

Duplication of *MdWRKY* Genes

Genes within a single genome can be classified as singletons, dispersed duplicates, proximal duplicates, tandem duplicates and segmental/WGD duplicates depending on their copy number and genomic distribution (Wang et al. 2012). Duplication events can result in a clustered occurrence of family members through tandem amplification, or a scattered occurrence through segmental duplication of chromosomal regions (De Grassi et al. 2008). Two or more homologous genes within a 100 kb range distance were defined as tandem duplicates. Twenty tandem duplication regions clustered with 44 *MdWRKY* genes, including 10 in Group I, 5 in Group IIa, 4 in Group IIb, 11 in Group IIc, 4 in Group IId, 2 in Group IIe and 8 in Group III, were identified on chr1, 3, 4, 6, 7, 8, 9, 11, 12 and 15 (Fig. 1, Table S1). Chromosomes 1 and chromosomes 12 had four clusters respectively, indicating hot spots of *WRKY* gene distribution. A total 78 *MdWRKY* genes forming of 85 pairs with collinear relationships were identified involving in segmental duplication (Fig. 1, Table S2). Ks values of the *MdWRKY* gene pairs ranged from 0.15 to 1.95. It was recognized that two genome-wide duplications, the paleoduplication event corresponding to the γ triplication (Ks ~1.6) and a recent WGD (Ks ~0.2), were occurred in the evolution of the apple genome (Velasco et al. 2010). We inferred that the segmental duplications, such as *MdWRKY41b* /*MdWRKY11e* (Ks ~1.78), *MdWRKY3b*/*MdWRKY72f* (Ks ~1.82), *MdWRKY4a*/*MdWRKY72f* (Ks ~1.66), *MdWRKY7d*/ *MdWRKY40c* (Ks ~1.68), *MdWRKY23a*/*MdWRKY7b* (Ks ~1.73), *MdWRKY40b*/*MdWRKY46b* (Ks ~1.82) and *MdWRKY48b*/*MdWRKY29a* (Ks ~1.95) may have arisen from the paleoduplication (~140 MYA); Many duplicated gene pairs had Ks values 0.15 ~ 0.4, suggesting that these duplications may have been derived from a recent WGD (30 ~ 45 MYA). Tandem duplications had produced *MdWRKY* gene clusters or hotspots, whereas segment duplications had produced many homologs of *MdWRKY* genes on different chromosomes, as indicated in Fig. 1. These results suggested that tandem and segmental duplications could be the major mechanism for expansion of WRKY family in apple.

Transcription Level of *MdWRKY* Genes in Response to *A. alternata* AP Infection

WRKY TFs play crucial roles in responses to biotic stress (Guo et al. 2014; Wang et al. 2014; Jiang et al. 2014; Dong et al. 2003). To investigate the expression of *MdWRKY* genes response to biotic stresses at the transcriptional level, we conducted transcriptome sequencing analysis using leaf samples inoculated by *A. alternata* AP. The expression profile of *MdWRKY* genes in response to the pathogen by transcriptome sequencing analysis was shown in the heatmap (Fig. 5) and Table S5. A total of 97 *MdWRKY* transcripts were detectable in apple infected leaves, among which 63 showed significant changes. The heatmap was divided into six clusters (Fig. 5). Cluster 1 contained 30 *MdWRKY* genes, which were all up-regulated at 72 h. Moreover, most of them showed significantly

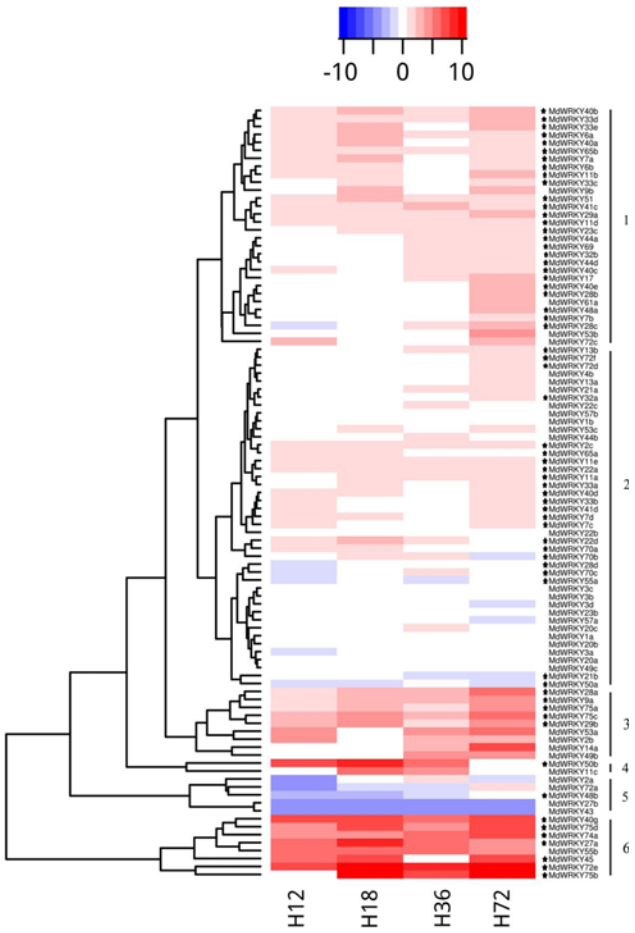

Fig. 5. Heat map of RNA-Seq expression of *MdWRKY* genes in response to *A. alternata* AP. Color scale of the dendrogram represents $\log_2$ Ratio value of treated sample to control sample. Asterisks indicate that the gene was significantly up or down regulated in the apple leaves (RPKM values≥ 1; $|\log_2 \text{Ratio}| \geq 1$).

differential expression except *MdWRKY9b*, *MdWRKY61a*, *MdWRKY53b* and *MdWRKY72c*. In clusters 2 (43 genes), a large number of genes were up-regulated, among which 25 were significantly induced, while 9 were down-regulated and 9 were not induced by pathogen infection. Cluster 3 and cluster 4 contained 9 and 2 genes respectively. The genes in cluster 3 were dramatically up-regulated at 36 h and 72 h, while in cluster 4 were up-regulated at 12 h or 18 h but not induced at 72 h by pathogen infection. Cluster 5 had 5 genes, which almost were down-regulated, but only one gene was significantly induced. In particular, cluster 6 contained 8 genes, which were highly dramatically up-regulated by pathogen infection at 12 h, 18 h, 36 h and 72 h. The RNA-seq data were also validated by qRT-PCR method. Thirteen *MdWRKY* genes with the significant differential expression induced by *A. alertnata* were selected for further analysis by real-time RT-PCR. As expected, real-time RT-PCR based expression patterns of these selected genes were consistent with those detected by RNA-seq analysis (Fig. S2).

Expression Profiles of *MdWRKY* Genes to Hormone Treatments

Plant hormones are indispensable roles in plant signaling networks (Bari and Jones 2009). For example, salicylic acid (SA), jasmonate (JA) and ethylene (Eth) are known to play important roles in regulating plant defense responses (Fujita et al. 2006). To explore the roles of *MdWRKY* genes under hormone treatments, we randomly selected 31 genes in the 63 *MdWRKY* genes, which showed differential expressions by pathogen infection based on the RNA-seq analysis, and detected their changes in transcript abundance under SA, MeJA and ethephon treatments by qRT-PCR (Fig. 6, Table S6). The 31 *MdWRKY* genes also showed differential expressions in response to at least one hormone treatment. Twenty-four *MdWRKY* genes showed differential expressions (>2-fold change) in response to SA treatment, whereas the other 7 showed no significant changes. Moreover, 5 *MdWRKY* genes (*MdWRKY29b*, *-51*, *-70b*, *-75c* and *-75d*) showed up-regulation (>2-fold change) at three time-points after the SA treatment. Twelve and 9 showed up- and down-regulation respectively at one or two time point(s), and no showed down-regulation at three time points. Twenty-one and 26 *MdWRKY* genes showed differential expressions in response to the MeJA and ethephon treatments, respectively. Two (*MdWRKY28c*, *-75d*) and five (*MdWRKY29b, -33e*, *-40b*, *-75c* and *-75d*) showed up-regulation upon the MeJA and ethephon treatments at three time points, respectively. Two (*MdWRKY48b*, *-55a*) were down regulated by the MeJA treatment at three time points, while no genes was down regulated by ethephon treatment at three time points. It is worth noticing that quite a few *MdWRKY* genes were induced by at least two hormone treatments except *MdWRKY23c*, *-11b*, *-28b* and *-22d*. Particularly, *MdWRKY75d* was dramatically up regulated by *A. alternata* infection also up regulated significantly by SA, MeJA and ethephon treatments. The highest transcript levels of *MdWRKY75d* increased 47.1, 6.6 and 82.4-fold after SA, MeJA and ethephon treatments. On the other hand, *MdWRKY48b*, which was down regulated by the fungus infection, was down regulated by SA, MeJA and ethephon treatments (Fig. 6, Table S6). The above results suggest that crosstalks exist in the hormone signaling pathways that mediate apple defense responses by the important regulators of WRKY TFs (Bari and Jones 2009).

Discussion

WRKYs are encoded by a multigene family. The number and composition of *WRKY* family members differ in various plants. In this study, a total of 119 *MdWRKY* genes in *Malus*

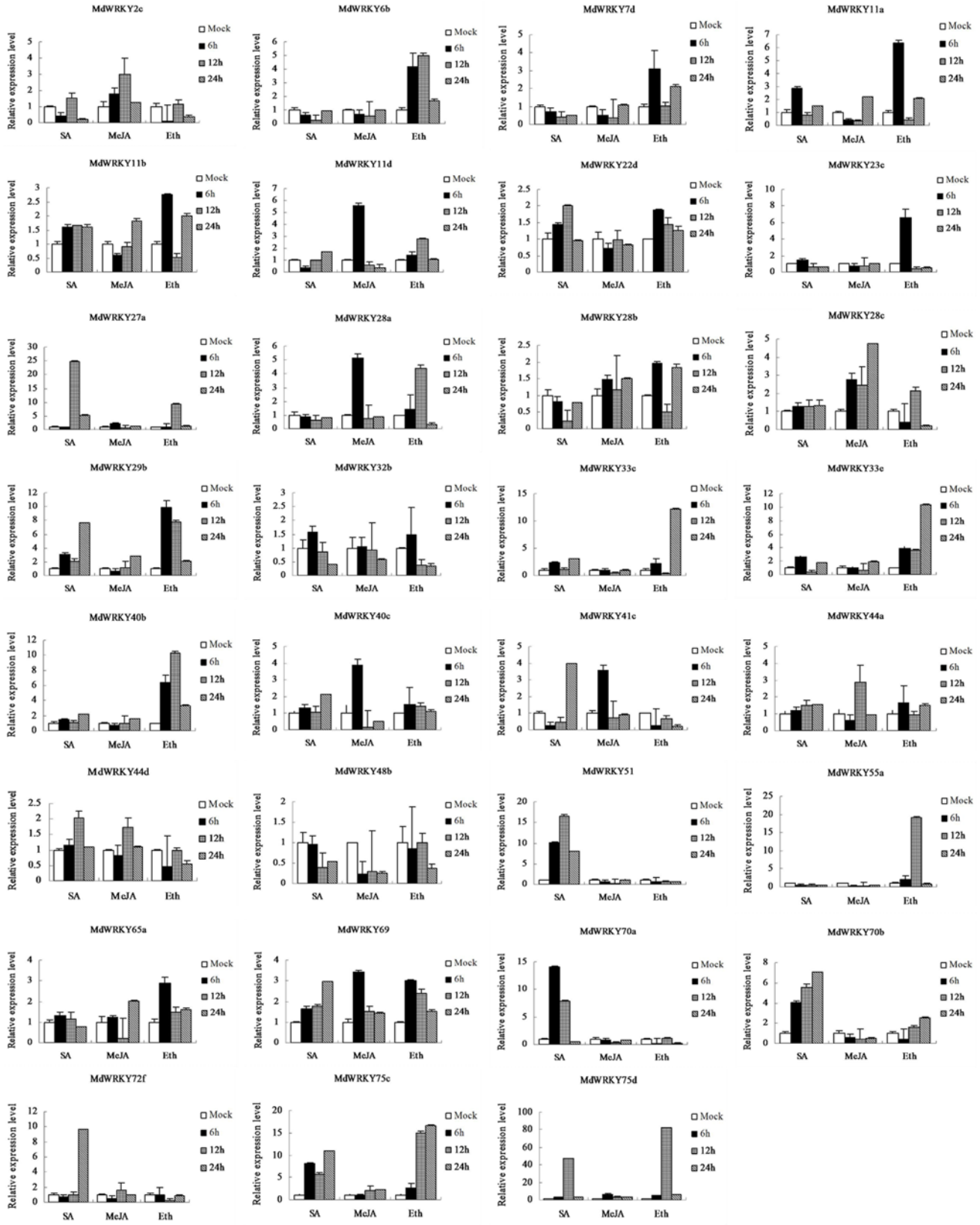

Fig. 6. Expression profiles of 31 *MdWRKY* genes in response to hormones. qRT-PCR expression analysis of *MdWRKY* genes in apple leaves treated by SA (A), MeJA (B) and Eth (C). Transcripts were normalized to Tubulin gene expression. The error bars of three biological replicates are presented.

domestica were identified through genome-wide analysis. The number is similar as 109 *PbWRKYs* identified in *pyrus bretschneideri* (Huang et al. 2015), but about 2 times as much as that in *Prunus persica* (58 *PpWRKYs*) (Chen et al. 2016) and in *Fragaria vesca* (62 *FvWRKYs*) (Wei et al. 2016), possibly because of the only one paleoduplication event that occurred in *P. persica* and *F. vesca,* whereas two genome replication events that occurred in *M. domestica* and *P. bretschneideri* during evolution. In a recent study, Meng et al. (2016) identified 127 WRKY family genes in *Malus domestica* genome. Compared with their report, eleven members were removed in our study due to their insufficient WRKY domain or redundancy, and three new members (MDP0000169621, MDP0000288378 and MDP0000200748) were identified in our report (Table S7).

The phylogenetic tree of *MdWRKY* genes indicates that Groups IIa and IIb, and Groups IId and IIe had close phylogenetic relationships respectively, and seemed to form monophyletic clades, that is, there are four major lineages within three groups of WRKY genes comprising Group IIa + Group IIb, Group IId + Group IIe, Group IIc + GroupI, and Group III. Gene duplication events, including tandem, segmental and whole genome duplications, play an important role not only in gene expansion and genomic rearrangement, but also in the diversification of gene function (Vision et al. 2000; Cannon et al. 2004). Recent genome-wide studies have revealed that the apple genomes experienced at least two genome duplications, one paleoduplication event and a recent WGD. Furthermore, a lot of studies proved that genome duplication contributed to the accumulation of gene families in apple, such as MAPK and MAPKK (Zhang et al. 2013), MYB (Cao et al. 2013), NAC (Su et al. 2013), LBD (Wang et al. 2013). In present study, 44 *MdWRKY* genes taken place 20 tandem duplication events and 78 *MdWRKY* genes taken place 85 segmental duplications were identified in apple genome (Fig. 1, Table S1-S2). The synteny analysis verified that the expansion of the WRKY gene family in apple was derived primarily from segmental duplications or WGD. This phenomenon also occurred in Arabidopsis (Cannon et al. 2004), grapevine (Wang et al. 2014), poplar (He et al. 2012) and pear (Huang et al. 2015).

Gene duplication generates functional redundancy (pseudogenization), subfunctionaliztion in which each daughter gene adopts part of the functions of their parental gene, and neofunctionalization (Zhang 2003). In apple, the paralogous pairs (*MdWRKY75a* and *MdWRKY75c*; *MdWRKY75a* and *MdWRKY75d*), which derived from segmental duplication, showed the same expression profiles during fungus infection. Three tandem duplication pairs (*MdWRKY33a* and *MdWRKY33b*, *MdWRKY44a* and *MdWRKY44d*, *MdWRKY40e* and *MdWRKY40c*) demonstrated the similar expression profiles, suggesting that some paralogs in apple are redundant. In contrast, the segmental duplications (*MdWRKY48b* and *MdWRKY23c*; *MdWRKY48b* and *MdWRKY29b*) and one tandem duplication pair (*MdWRKY70a* and *MdWRKY70c*) showed the different expression in response to pathogen, suggesting the expansion of the WRKY family followed by functional divergence might be to remove their redundancy (Yao et al. 2015).

The plant's innate immune system consists of two interconnected branches termed PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI) that initiate massive transcriptional reprogramming (Zhang et al. 2013). Extensive studies have established that plant WRKY TFs play important roles in the two branches of the plant innate immune system (Journot-Catalino et al. 2006; Eulgem and Somssich 2007). To know *MdWRKY* genes responding to pathogen, we performed transcriptomic analysis of apple leaves infected by *A. alternata* AP. The pathogen caused apple alternaria blotch disease which is one of the most serious fungal diseases affecting apple around the world especially in East Asia (Saito et al. 2001). In the experiment, more than half (63 out of 119) of *MdWRKY* genes were responsive to *A. alternata* AP infection. Except *MdWRKY48b, MdWRKY55a* and *MdWRKY70b* which showed down-expression at certain time points, most of *MdWRKY* genes showed up regulation during the infection phase. The result was similar with that most of *AtWRKY* genes were up-regulated in Arabidopsis upon pathogen infection (Dong et al. 2003). It indicated that a large numbers of *WRKY* genes in apple are associated with plant defense responses, meanwhile, a number of *MdWRKY* genes were either not expressed, or not altered significantly in pathogen-infected plants, suggesting that some of the *MdWRKY* genes might be involved in other biological processes.

SA, JA or Eth functions as signaling substances, and its accumulation plays a central role in the activation of defense reactions leading to plant disease resistance (Delaney et al. 1994; Bari and Jones 2009). WRKY TFs are important regulators of SA-dependent defense responses (Wang et al. 2006; Eulgem and Somssich 2007). In Arabidopsis, nearly 70% *AtWRKYs* are differentially regulated by pathogen infection or salicylic acid (SA) treatment (Dong et al. 2003). Moreover, some WRKY TFs have been shown to involve in the crosstalk between SA, JA/Eth-mediated defense response signaling systems. For example, *AtWRKY70* integrates signals from mutually antagonistic SA- and JA-dependent pathways in plant defense responses (Li 2004). In banana, SA and MeJA treatments can activate *PR* and *WRKY* gene expressions, and WRKY TFs binding to PR promoters may be attributed to SA- and MeJA-induced pathogen resistance (Tang et al. 2013). In our experiment, the 31 chosen *MdWRKY* genes, including 6 from Group I, 21 from Group II and 4 from Group III, which showed differential expressions by pathogen

infection also showed differential expressions in response to SA, MeJA or ethephon treatments. There are 24, 21 and 26 *MdWRKY* genes showed differential expression upon SA, MeJA and ethephon treatment, respectively. Among the 24 SA-regulated WRKY genes, 17 were also regulated by MeJA and 20 regulated by ethephon. *MdWRKY75c*, *MdWRKY75d* and *MdWRKY28a*, the orthologous genes of *AtWRKY75* and *AtWRKY28* in Arabidopsis, displayed high expression levels after infection with *A. alternata* AP (Fig. 2, Fig. 5). In Arabidopsis, over expressing *AtWRKY75* and *AtWRKY28* induced oxidative burst in host plants and consequently triggered plant resistance to infection by *S. sclerotiorum*, which is most probably associated with the activation of the SA- and JA/Eth-mediated defense signaling pathways (Chen et al. 2013). Similarly, *MdWRKY75c* and *MdWRKY75d* also drastically activated by SA, MeJA and ethephon treatments, and *MdWRKY28a* activated by MeJA/ethephon treatments. These results suggested that the regulation of MdWRKY TFs might be mediated by the crosstalk between hormone signaling pathways in apple defense against pathogen infection.

Conclusion

In summary, 119 *MdWRKY* genes were identified in apple genome. These *WRKY* genes were anchored on 17 chromosomes in apple with uneven distribution, and were classified into three main groups. The members in same groups and subgroups contain similar exon–intron structures and motif compositions. The expression profiles of *MdWRKY* genes under hormone treatments and inoculation by *A. alternata* AP suggested that a large numbers of *MdWRKY* genes play roles in plant defense response regulation and respond to hormone treatments and the crosstalks exist between SA- and JA/Eth-mediated defense response.

Material and Methods

Identification of WRKY Genes in Apple Genome

The complete genome, proteome sequences and GFF (General Feature Format) of Arabidopsis and apple (Malus domestica) were obtained from TAIR (The Arabidopsis Information Resource version 10; http://www.arabidopsis.org) and GDR (Genome Database for Rosaceae; http://www.rosaceae.org/) and Phytozome v10 (http://www. phytozome.net), respectively. In proteome datasets, if two or more protein sequences at the same locus were identical where they overlapped, the longest sequences were selected. We use an HMMER-BLASTp-InterProScan strategy to search for WRKY genes in genome. A HMM profile for the WRKY domain (PF03106) was downloaded from Pfam (http://pfam.janelia.org/). HMMER (Eddy 1998) was used to search a customized database containing the protemome of each species with the threshold set at of the Pfam GA gathering cutoff. The HMMER selected proteins were used for a BLASTp query of the original protein database. Finally, the BLASTp hits were scanned for WRKY domains using InterPro9 WRKY genes.

Phylogenetic Construction of *MdWRKYs*

Multiple sequence alignment using MUSCLE (Edgar 2004) with the default parameters was performed with the WRKY domain of protein sequences. The phylogenetic trees using multiple sequence alignment of domain sequences were constructed using the Neighbor-Joining (NJ). NJ trees were constructed using MEGA 6.0 software with a bootstrap test of 1000 replicates.

Chromosomal Locations, Synteny Analysis, Conserved Motifs and Exon–intron Structures of Apple *WRKY* Genes

The genes were plotted separately onto the chromosomes according to gene location in the chromosome in the GFF file using a programmed Perl script. Genes in a range of 100 kb distance were considered to be tandem duplicates. The microsynteny between each pair of members was detected by using the MicroSyn software (Cai et al. 2011). The parameters were set as follows: window size of 50 genes, tandem gap value of 2, expected threshold value cut off of 0.01, and 3 homologous pairs to define a syntenic segment. The conserved motifs in the WRKY proteins were detected by MEME (http://meme.nbcr.net/meme/cgibin/meme.cgi), with the following parameters: number of repetitions: any; maximum number of motifs: 50; and the optimum motif widths: 6–200 amino acid residues. The exon–intron structures of *MdWRKY* genes were predicted by comparing coding sequences with their corresponding genomic sequences using the online GSDS program (http://gsds.cbi.pku.edu.cn).

Plant Material and Treatments

Plant Material

Plant materials were taken from three-year-old apple plants (*Malus ×domestica* Borkh. cv. 'Starking Delicious') grafted on *Malus robusta* Rehd. stocks which were grown in the greenhouse in Nanjing Agricultural University, located in Nanjing, Jiangsu Province, China.

Fungus Infection and Hormone Treatments

The inoculation method was performed according to the protocol of Abe et al. (2010). *A. alternata* AP was expanded on potato dextrose agar (PDA) medium at 26°C under dark conditions for 5 d. The forth and the fifth youngest opened leaves from the shoot tips were inoculated by mycelia biomasses of *A. alternata* AP. For mock inoculations, PDA medium cake was put on instead of the mycelia biomasses. Leaves sampled at 0, 12, 18, 36, and 72 hours post-inoculation (HPI) were used to extract RNA.

For hormone treatments, apple leaves were subjected to 5 mM Salicylic acid (SA), 50 μM methyl jasmonate (MeJA) and 5 mM ethephon. Leaves samples were then collected at 6 h, 12 h and 24 h intervals, respectively. Leaves sprayed with sterile water and similarly harvested were used as a control. Each sample contained three parallel leaves that represented three biological replicates.

Library Construction and RNA-sequencing Analysis

Total RNA was isolated using a cetyltrimethyl ammonium bromide (CTAB) method (Pavy et al. 2008; Chang et al. 1993). Equal quantities of RNA from three biological replications for each stage were pooled for constructing a cDNA library. Oligo-(dT) magnetic beads were used to isolate poly-(A) mRNA from total RNA, and mRNA was fragmented

in fragmentation buffer. Using these short fragments (≈200 bp) as templates, random hexamer-primers were used to synthesize first-strand cDNA. Second-strand cDNA was synthesized using buffer, dNTPs, RNaseH and DNA polymerase I. Short double-stranded cDNA were purified with QiaQuick PCR extraction kit (Qiagen, Venlo, Netherlands), resolved with EB buffer for end reparation and for adding poly (A), then ligated to sequencing adapters. After purification via agarose gel electrophoresis, suitable fragments were enriched by PCR amplification. Finally, the libraries were sequenced via the Illumina HiSeqTM2000 platform at the Beijing Genomics Institute (Shenzhen, China; http://www.genomics.cn/index.php), following the manufacturer's protocols.

Raw reads from the image data output were generated by Base Calling. Clean reads were generated by removing adaptors, low-quality reads, and aligned to reference sequences using SOAPaligner/soap2 (Li et al. 2009). The gene expression levels were calculated by using the RPKM (Reads Per kb per Million reads) method (Mortazavi et al. 2008). Differential expressed genes were analyzed as described previously (Audic and Claverie 1997). Additionally, the FDR (False Discovery Rate) ≤ 0.001 and the absolute value of log2 | Ratio | ≥ 1 were used as the threshold to judge the significance of gene expression differences.

Quantitative RT-PCR analysis

Total RNA was isolated as above. Then RNA was treated with DNase I (RNase-free DNase set, Qiagen, Hilden, Germany) to eliminate trace of DNA, and RNA quality and concentration were measured using NanoDrop 1000 (Thermo Scientific, Wilmington, DE). First strand cDNA was synthesized from 1 μg total RNA using cDNA reverse transcription kit (Applied Biosystems, Foster city, CA) with RNase Inhibitor (RNase out, Invitrogen Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. Quantitative RT-PCR was conducted using SYBR green (TaKaRa Biotechnology) on an ABI 7300 Real Time PCR System (Applied Biosystems, Foster City, CA USA). Each reaction mixture contained 10.0 μL SYBR Premix ExTaq (Tli RNaseH Plus) (TaKaRa Biotechnology), 1.0 μL cDNA template, 0.4 μL each primer and 8.2 μL sterile distilled H_2O. Each reaction was performed in three technical replicates. Specific primers used for qRT-PCR analysis are shown in Supplementary Table S4. Transcripts were normalized to *Mdtubulin* (GenBank accession number AJ421411) gene expression. The parameters of Cycling were 95°C for 3 min, 40 cycles at 95°C for 20 s, 60°C for 20 s, and 72°C for 40 s. The parameters of the melt-curve analyses were performed with 95°C for 15 s and then a constant increase from 60°C to 95°C. $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001) was used to calculate the relative expression of *MdWRKY* genes.

Acknowledgments

This work was supported by the National Science foundation of China (31171935), (31560551), the fundamental Research Funds for the Central Universities (KYZ201310) and Jiangsu Agricultural Science and Technology Innovation fund [CX(15)1022]. We are also grateful to Prof.Guo Yunzhong in Northwest A&F University, China for providing *A. alternata* AP stain.

Author Contributions

Shuai Liu carried out a considerable part of the experiments, analyzed the data, and wrote the paper. Changguo Luo, Renhe Sha, Shenchun Qu, Binhua Cai were involved in the experimental process and in the discussion of the obtained results. Sanhong Wang designed the experiment and proofread the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

Supporting Information

Fig. S1. Multiple sequence alignment of the WRKY domain among *MdWRKY* genes.
Fig. S2. Validation of RNA-Seq analysis by qRT-PCR. RPKM (reads per kilobase per million) values obtained with RNA-Seq. qRT-PCR expression analysis of 13 *MdWRKY* genes in apple leaves infected with *A. alternata* AP. Transcripts were normalized to tubulin gene expression. The error bars of three biological replicates are presented.
Table S1. Tandem duplication events in the *MdWRKY* genes.
Table S2. Segmental duplication genes in the *MdWRKY* genes.
Table S3. Motif sequences of HARF, LxLxL and LxxLL.
Table S4. List of qRT-PCR primers for selected *MdWRKY* genes.
Table S5. The expression level of *MdWRKY* genes in RNA-seq.
Table S6. *MdWRKY* genes expression patterns determined by RNA-seq and qRT-PCR.
Table S7. Comparison of the gene names and the gene IDs of putative MdWRKYs in the study and Meng et al. (2016).

References

Abe K, Iwanami H, Kotoda N, Moriya S, Takahashi Sumiyoshi S (2010) Evaluation of apple genotypes and *Malus* species for resistance to Alternaria blotch caused by *Alternaria alternata* apple pathotype using detached-leaf method. Plant Breeding 129:208–218

Audic S, Claverie JM (1997) The significance of digital gene expression profiles. Genome Res 7:986–995

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208

Bari R, Jones JDG (2009) Role of plant hormones in plant defence responses. Plant Mol Biol 69:473–488

Bulajic A, Filajdic N, Babovic M, Sutton TB (1996) First report of *Alternaria mali* on apples in Yugoslavia. Plant Dis 80:709

Cai B, Yang X, Tuskan GA, Cheng ZM (2011) MicroSyn: a user friendly tool for detection of microsynteny in a gene family. BMC Bioinformatics 12:79

Cai C, Niu E, Du H, Zhao L, Feng Y, Guo W (2014) Genome-wide analysis of the WRKY transcription factor gene family in *Gossypium raimondii* and the expression of orthologs in cultivated tetraploid cotton. The Crop Journal 2:87–101

Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol 4:10

Cao Z, Zhang S, Wang R, Zhang R, Hao Y (2013) Genome wide analysis of the apple MYB transcription factor family allows the identification of *MdoMYB121* gene confering abiotic stress tolerance in plants. PLoS ONE 8:e69955

Chang S, Jeff P, John C (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Rep 11:113–116
Chen C, Chen Z (2002) Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced Arabidopsis transcription factor. Plant Physiol 129:706–716
Chen C, Chen Z (2000) Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco. Plant Mol Biol 42:387–396
Chen M, Tan Q, Sun M, Li D, Fu X, Chen X, Xiao W, Li L, Gao D (2016) Genome-wide identification of WRKY family genes in peach and analysis of WRKY expression during bud dormancy. Mol Genet Genomics 291(3):1319-1332
Chen X, Liu J, Lin G, Wang A, Wang Z, Lu G (2013) Overexpression of *AtWRKY28* and *AtWRKY75* in Arabidopsis enhances resistance to oxalic acid and *Sclerotinia sclerotiorum*. Plant Cell Rep 32:1589–1599
Choi C, Park YH, Kwon SI, Yun C, Ahn I, Park SR, Hwang D (2014) Identification of AtWRKY75 as a transcriptional regulator in the defense response to *Pcc* through the screening of Arabidopsis activation-tagged lines. Plant Biotechnol Rep 8:183–192
Ciolkowski I, Wanke D, Birkenbihl RP, Somssich IE (2008) Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. Plant Mol Biol 68:81–92
Dang F, Wang Y, She J, Lei Y, Liu Z, Eulgem T, Lai Y, Lin J, Yu L, Lei D, Others (2014) Overexpression of *CaWRKY27*, a subgroup IIe WRKY transcription factor of *Capsicum annuum*, positively regulates tobacco resistance to *Ralstonia solanacearum* infection. Physiol Plantarum 150:397–411
Dang F, Wang Y, Yu L, Eulgem T, Lai Y, Liu Z, Wang X, Qiu A, Zhang T, Lin J, Chen Y, Guan D, Cai H, Mou S, He S (2013) CaWRKY40, a WRKY protein of pepper, plays an important role in the regulation of tolerance to heat stress and resistance to *Ralstonia solanacearum* infection. Plant, Cell Environ 36:757–774
De Grassi A, Lanave C, Saccone C (2008) Genome duplication and gene-family evolution: the case of three OXPHOS gene families. Gene 421:1–6
Delaney TP, Uknes S, Vernooij B, Friedrich L, Weymann K, Negrotto D, Gaffney T, Gut-Rella M, Kessmann H, Ward E, Ryals J (1994) A central role of salicylic acid in plant disease resistance. Science 266:1247–1250
Dong J, Chen C, Chen Z (2003) Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. Plant Mol Biol 51:21–37
Duan GF, Li L J, Liu QL (2014). A wrky transcription factor from malus domestica, negatively regulates dehydration stress in transgenic arabidopsis. Acta Physiol Plant, 36(2):541-548
Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14:755–763
Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113
Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends Plant Sci 5:199–206
Eulgem T, Rushton PJ, Schmelzer E, Hahlbrock K, Somssich IE (1999) Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. EMBO J 18:4689–4699
Eulgem T, Somssich IE (2007) Networks of WRKY transcription factors in defense signaling. Curr Opin Plant Biol 10:366–371
Fan H, Wang F, Gao H, Wang L, Xu J, Zhao Z (2011) Pathogen-induced MdWRKY1 in 'Qinguan' apple enhances disease resistance. J Plant Biol 54:150–158
Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K (2006) Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Curr Opin Plant Biol 9:436–442
Guo C, Guo R, Xu X, Gao M, Li X, Song J, Zheng Y, Wang X (2014) Evolution and expression analysis of the grape (*Vitis vinifera* L.) *WRKY* gene family. J Exp Bot 65:1513–1528
He H, Dong Q, Shao Y, Jiang H, Zhu S, Cheng B, Xiang Y (2012) Genome-wide survey and characterization of the *WRKY* gene family in *Populus trichocarpa*. Plant Cell Rep 31:1199–1217
Huang X, Li K, Xu X, Yao Z, Jin C, Zhang S (2015) Genome-wide analysis of WRKY transcription factors in white pear (*Pyrus bretschneideri*) reveals evolution and patterns under drought stress. BMC Genomics 16
Ishiguro S, Nakamura K (1994) Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5'-upstream regions of genes coding for sporamin and β-amylase from sweet potato. Mol General Genet 244:563–571
Jiang Y, Duan Y, Yin J, Ye S, Zhu J, Zhang F, Lu W, Fan D, Luo K (2014) Genome-wide identification and characterization of the *Populus* WRKY transcription factor family and analysis of their expression in response to biotic and abiotic stresses. J Exp Bot 65:6629–6644
Journot-Catalino N, Somssich IE, Roby D, Kroj T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis thaliana*. Plant Cell 18:3289–3302
Li J (2004) The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell 16:319–331
Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J (2009) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25:1966–1967
Li S, Fu Q, Chen L, Huang W, Yu D (2011) *Arabidopsis thaliana* WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. Planta 233:1237–1252
Ling J, Jiang W, Zhang Y, Yu H, Mao Z, Gu X, Huang S, Xie B (2011) Genome-wide analysis of WRKY gene family in *Cucumis sativus*. BMC Genomics 12:471
Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method. Methods 25:402–408
Mangelsen E, Kilian J, Berendzen KW, Kolukisaoglu UH, Harter K, Jansson C, Wanke D (2008) Phylogenetic and comparative gene expression analysis of barley (*Hordeum vulgare*) WRKY transcription factor family reveals putatively retained functions between monocots and dicots. BMC Genomics 9:194
Meng D, Li Y, Bai Y, Li M, Cheng L (2016) Genome-wide identification and characterization of WRKY transcriptional factor family in apple and analysis of their responses to waterlogging and drought stress. Plant Physiol Bioch 103:71–83
Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628
Murray SL, Ingle RA, Petersen LN, Denby KJ (2007) Basal resistance against Pseudomonas syringae in Arabidopsis involves WRKY53 and a protein with homology to a nematode resistance protein. Mol Plant Microbe Interact 20:1431–1438
Pavy N, Boyle B, Nelson C, Paule C, Giguere I, Caron S, Parsons LS, Dallaire N, Bedon F, Berube H, Cooke J, Mackay J (2008) Identification of conserved core xylem gene sets: conifer cDNA microarray development, transcript profiling and computational analyses. New Phytol 180:766–786
Qin Y, Tian Y, Han L, Yang X (2013) Constitutive expression of a salinity-induced wheat WRKY transcription factor enhances

salinity and ionic stress tolerance in transgenic *Arabidopsis thaliana*. Bioch Biophy Res Commun 441:476–481
Rairdan G, Moffett P (2007) Brothers in arms? Common and contrasting themes in pathogen perception by plant NB-LRR and animal NACHT-LRR proteins. Microbes Infect 9:677–686
Ramamoorthy R, Jiang SY, Kumar N, Venkatesh PN, Ramachandran S (2008) A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. Plant Cell Physiol 49:865–879
Ramiro D, Jalloul A, Petitot A, Grossi De Sá MF, Maluf MP, Fernandez D (2010) Identification of coffee WRKY transcription factor genes and expression profiling in resistance responses to pathogens. Tree Genet Genomes 6:767–781
Rinerson CI, Rabara RC, Tripathi P, Shen QJ, Rushton PJ (2015) The evolution of WRKY transcription factors. BMC Plant Biol 15:66
Robatzek S, Somssich IE (2001) A new member of the Arabidopsis WRKY transcription factor family, *AtWRKY6*, is associated with both senescence- and defence-related processes. Plant J 28:123–133
Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. Trends in Plant Sci 15:247–258
Ryu HS, Han M, Lee SK, Cho JI, Ryoo N, Heu S, Lee YH, Bhoo SH, Wang GL, Hahn TR, Jeon JS (2006) A comprehensive expression analysis of the *WRKY* gene superfamily in rice plants during defense response. Plant Cell Rep 25:836–847
Saito A, Nakazawa N, Suzuki M (2001) Selection of mutants resistant to Alternaria blotch from in vitro-cultured apple shoots irradiated with X-and γ-rays. J Plant Physiol 158:391–400
Su H, Zhang S, Yuan X, Chen C, Wang X, Hao Y (2013) Genome-wide analysis and identification of stress-responsive genes of the NAM--ATAF1, 2--CUC2 transcription factor family in apple. Plant Physiol Bioch 71:11–21
Sun C, Palmqvist S, Olsson H, Bor E N M, Ahlandsberg S, Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the iso1 promoter. Plant Cell 15:2076–2092
Tang Y, Kuang J, Wang F, Chen L, Hong K, Xiao Y, Xie H, Lu W, Chen J (2013) Molecular characterization of *PR* and *WRKY* genes during SA- and MeJA-induced resistance against *Colletotrichum musae* in banana fruit. Postharvest Biol Technol 79:62–68
Tiwari SB, Hagen G, Guilfoyle TJ (2004) Aux/IAA proteins contain a potent transcriptional repression domain. Plant Cell 16:533–543
Tripathi P, Rabara RC, Rushton PJ (2014) A systems biology perspective on the role of WRKY transcription factors in drought responses in plants. Planta 239:255–266
Ülker B, Mukhtar MS, Somssich IE (2007) The WRKY70 transcription factor of Arabidopsis influences both the plant senescence and defense signaling pathways. Planta 226:125-137
Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagné D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouzé P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nat Genet 42:833–839
Vision TJ, Brown DG, Tanksley SD (2000) The origins of genomic duplications in Arabidopsis. Science 290:2114–2117
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:e123
Wang L, Zhu W, Fang L, Sun X, Su L, Liang Z, Wang N, Londo JP, Li S, Xin H (2014) Genome-wide identification of *WRKY* family genes and their response to cold stress in *Vitis vinifera*. BMC Plant Biol 14:103
Wang M, Vannozzi A, Wang G, Liang Y, Tornielli GB, Zenoni S, Cavallini E, Pezzotti M, Cheng ZM (2014) Genome and transcriptome analysis of the grapevine (*Vitis vinifera* L.) WRKY gene family. Hort Res 1:14016
Wang X, Zhang S, Su L, Liu X, Hao Y (2013) A genome-wide analysis of the LBD (LATERAL ORGAN BOUNDARIES domain) gene family in *Malus domestica* with a functional characterization of *MdLBD11*. PloS ONE 8:e57044
Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40:e49
Wei W, Hu Y, Han YT, Zhang K, Zhao FL, Feng JY (2016) The WRKY transcription factors in the diploid woodland strawberry *Fragaria vesca*: Identification and expression analysis under biotic and abiotic stresses. Plant Physiol Biochem 105:129–144
Wu KL, Guo ZJ, Wang HH, Li J (2005) The WRKY family of transcription factors in rice and Arabidopsis and their origins. DNA Res 12:9–26
Yao Q, Xia E, Liu F, Gao L (2015) Genome-wide identification and comparative expression analysis reveal a rapid expansion and functional divergence of duplicated genes in the WRKY gene family of cabbage, *Brassica oleracea* var. *capitata*. Gene 557:35–42
Zhang J (2003) Evolution by gene duplication: an update. Trends Ecol Evol 18:292–298
Zhang S, Xu R, Luo X, Jiang Z, Shu H (2013) Genome-wide identification and expression analysis of MAPK and MAPKK gene family in *Malus domestica*. Gene 531:377–387
Zhang Y, Lubberstedt T, Xu M (2013) The genetic and molecular basis of plant resistance to pathogens. J Genet Genomics 40:23–35
Zhang Y, Wang L (2005) The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol Biol 5:1
Zheng Z, Mosher SL, Fan B, Klessig DF, Chen Z (2007) Functional analysis of Arabidopsis WRKY25 transcription factor in plant defense against Pseudomonas syringae. BMC Plant Biol 7:2
Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J 48:592-605
Zhou Q, Tian A, Zou H, Xie Z, Lei G, Huang J, Wang C, Wang H, Zhang J, Chen S (2008) Soybean WRKY-type transcription factor genes, *GmWRKY13*, *GmWRKY21*, and *GmWRKY54*, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants. Plant Biotechnol 6:486–503
Zou C, Jiang W, Yu D (2010) Male gametophyte-specific WRKY34 transcription factor mediates cold sensitivity of mature pollen in Arabidopsis. J Exp Bot 61:3901–3914