Genome-wide characterization and expression analyses of the *MYB* superfamily genes during developmental stages in Chinese jujube

Ji Qing [1], **Wang Dawei** [1], **Zhou Jun** [Corresp., 1, 2], **Xu Yulan** [1], **Shen Bingqi** [1], **Zhou Fan** [1]

[1] Southwest Forestry University, Key Laboratory for Forest Resource Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Kunming, Yunnan, China
[2] North Minzu University, College of Life Science and Engineering, Yinchuan, China

Corresponding Author: Zhou Jun
Email address: zhoujunbo@163.com

The MYB transcription factor superfamily, one of the largest gene superfamilies, regulates a variety of physiological processes in plants. Although many MYB superfamily genes have been identified in plants, the MYB TFs in Chinese jujube (*Ziziphus jujuba* Mill.) have not been fully identified and characterized. Additionally, the functions of these genes remain unclear. In total, we identified 171 *MYB* superfamily genes in jujube and divided them into 5 subfamilies containing 99 genes of the R2R3-MYB subfamily, 58 genes of the MYB-related subfamily, 4 genes of the R1R2R3-MYB subfamily, 1 gene of the 4R-MYB subfamily, and 9 genes of the atypical MYB subfamily. The 99 R2R3-MYB genes of jujube were divided into 35 groups, C1 to C35, and the 58 MYB-related genes were divided into the following groups: the R-R-type, CCA1-like, I-box-binding-like, TBP-like, CPC-like, and Chinese jujube-specific groups. *ZjMYB* genes in jujube were well supported by additional highly conserved motifs and exon/intron structures. Most R1 repeats of MYB-related proteins comprised the R2 repeat and had highly conserved EED and EEE residue groups in jujube. Three tandem duplicated gene pairs were found on twelve chromosomes in jujube. According to an expression analysis of 126 *ZjMYB* genes, MYB-related genes played important roles in jujube development and fruit-related biological processes. The total flavonoid content of jujube fruit decreased as ripening progressed. Ninety-three expressed genes were identified in the RNA-sequencing data from jujube fruit, and 56 *ZjMYB* members presented significant correlations with total flavonoid contents by correlation analysis. Five pairs of paralogous *MYB* genes within jujube were composed of 9 jujube *MYB* genes. Fourteen *ZjMYB* genes had the same homology to the *MYB* genes of *Arabidopsis* and peach, indicating that these 14 *MYB* genes and their orthologs probably existed before the ancestral divergence of the MYB superfamily. We used a synteny analysis of *MYB* genes between jujube and *Arabidopsis* to predict that the functions of the *ZjMYBs* involve

flavonoid/phenylpropanoid metabolism, the light signaling pathway, auxin signal transduction, and responses to various abiotic stresses (cold, drought and salt stresses). Additionally, we speculate that *ZjMYB108* is an important transcription factor involved in the flavonoid metabolic pathway. This comprehensive analysis of *MYB* superfamily genes in jujube lay a solid foundation for future comprehensive analyses of *ZjMYB* gene functions.

Genome-wide characterization and expression analyses of the *MYB* superfamily genes during developmental stages in Chinese jujube

Ji Qing [1], Wang Dawei [1], Zhou Jun [1,2], Xu Yulan [1], Shen Bingqi [1], Zhou Fan [1]

[1] Key Laboratory for Forest Resource Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming, Yunnan, China

[2] College of Life Science and Engineering, North Minzu University, Yinchuan, China

Corresponding Author:

Zhou Jun [1,2]

NO. 300, Bailong Temple, Panlong District, Kunming, 650224, Yunnan, China

Email address: zhoujunbo@163.com

Abstract

The MYB transcription factor superfamily, one of the largest gene superfamilies, regulates a variety of physiological processes in plants. Although many MYB superfamily genes have been identified in plants, the MYB TFs in Chinese jujube (*Ziziphus jujuba* Mill.) have not been fully identified and characterized. Additionally, the functions of these genes remain unclear. In total, we identified 171 *MYB* superfamily genes in jujube and divided them into 5 subfamilies containing 99 genes of the R2R3-MYB subfamily, 58 genes of the MYB-related subfamily, 4 genes of the R1R2R3-MYB subfamily, 1 gene of the 4R-MYB subfamily, and 9 genes of the atypical MYB subfamily. The 99 R2R3-MYB genes of jujube were divided into 35 groups, C1 to C35, and the 58 MYB-related genes were divided into the following groups: the R-R-type, CCA1-like, I-box-binding-like, TBP-like, CPC-like, and Chinese jujube-specific groups. *ZjMYB* genes in jujube were well supported by additional highly conserved motifs and exon/intron structures. Most R1 repeats of MYB-related proteins comprised the R2 repeat and had highly conserved EED and EEE residue groups in jujube. Three tandem duplicated gene pairs were found on twelve chromosomes in jujube. According to an expression analysis of 126 *ZjMYB* genes, MYB-related genes played important roles in jujube development and fruit-related biological processes. The total flavonoid content of jujube fruit decreased as ripening progressed. Ninety-three expressed genes were identified in the RNA-sequencing data from jujube fruit, and 56 *ZjMYB* members presented significant correlations with total flavonoid contents by correlation analysis. Five pairs of paralogous *MYB* genes within jujube were composed of 9 jujube *MYB* genes. Fourteen *ZjMYB* genes had the same homology to the *MYB* genes of *Arabidopsis* and peach, indicating that these 14 *MYB* genes and their orthologs probably existed before the ancestral divergence of the MYB superfamily. We used a synteny analysis of *MYB* genes between jujube and *Arabidopsis* to predict that the functions of the *ZjMYBs* involve flavonoid/phenylpropanoid metabolism, the light signaling pathway, auxin signal transduction, and responses to various abiotic stresses (cold, drought and salt stresses). Additionally, we speculate that *ZjMYB108* is an important transcription factor involved in the flavonoid metabolic pathway. This comprehensive analysis of *MYB*

superfamily genes in jujube lay a solid foundation for future comprehensive analyses of *ZjMYB* gene functions.

Introduction

Transcription factors (TFs) play essential roles in plants by controlling the expression of genes, activating or inhibiting the transcription of other genes, or interacting with other TFs to regulate gene transcription (Singh, Foley & Onate-Sanchez, 2002; Liu, White & MacRae, 1999). The MYB family is a large TF family present in all eukaryotes, and MYBs regulate a variety of physiological processes in plants (Riechmann et al., 2000). The first identified MYB gene was *v-myb* in 1982 from the avian myeloblastosis virus, a chicken oncogene that leads to acute myeloblastic leukemia (Klempnauer et al., 1982), thus leading to the name *myb*. C-*myb,* a v-MYB-related gene was subsequently found in animal cells. Corresponding genes were also identified in human tumor cells (A-MYB and B-MYB) at the same time and these genes were found to modulate cell proliferation, differentiation, and apoptosis (Weston, 1998). The first identified *MYB* gene of a plant was C1, which was isolated from *Zea mays* (Paz-Ares et al., 1987) and is involved in anthocyanin biosynthesis. As the *Arabidopsis MYB* gene family has gradually been identified (Stracke, Werber & Weisshaar, 2001; Chen et al., 2006), more *MYB* genes have been identified in many other plants (Zhang & Ma et al., 2018; Li X et al., 2016; Hou et al., 2014; Zhou et al., 2015); these genes are identified by the family-specific feature of a highly conserved MYB domain at the N-terminus (Lipsick, 1996; Mmadi et al., 2017; Dubos et al., 2010). This domain gene usually comprises 1-4 imperfect amino acid sequence repeats (R1-R4) of approximately 50-53 amino acids (Dubos et al., 2010; Mmadi et al., 2017), with each forming three α-helices, and the second and third helices form a helix-turn-helix (HTH) motif (Ogata et al., 1996; Dubos et al., 2010). *MYB* genes are classified into the following subfamilies according to the number of MYB imperfect tandem repeats (Rs) of the proteins: MYB-related (or 1R-MYB, one R), R2R3-MYB (two Rs), R1R2R3-MYB (three Rs), and 4R-MYB (four Rs) (Dubos et al., 2010; Zhang & Ma et al., 2018).

Many studies have shown that MYB TFs are involved in physiological and biochemical processes

in plants and responses to various biotic as well as abiotic stresses (Abe et al., 2003; Agarwal et al., 2006; Andrew et al., 2007; Cominelli & Tonelli, 2009; Raffaele et al., 2008), and the function of MYB has also been studied in detail in some MYB proteins. For example, *MYB7* plays a role in kiwifruit (*Actinidia chinensis*) through transcriptional activation of metabolic pathway genes to modulate carotenoid and Chl pigment accumulation in tissues (Ampomah--Dwamena et al., 2018). The *AtDIV2* of the R-R-type *MYB* gene in *Arabidopsis* is required for ABA signaling and plays a negative role in salt stress (Fang et al., 2018). In tobacco, overexpression of the *PbrMYB5* gene enhanced tolerance to chilling stresses (Xing et al., 2018). In tomato (*Solanum lycopersicum* L.) plants, the R3-MYB gene is involved in a feedback mechanism, and if it is activated by endogenous or exogenous stimuli, anthocyanin production is inhibited (Colanero, Perata & Gonzali, 2018). During continuous light treatments, the levels of *McMYB10* increased and promoted the expression levels of *McCOP1-1* and *McCOP1-2* as well as anthocyanin biosynthesis in crabapple (Li et al., 2018). MYB TFs are also involved in flavonoid/phenylpropanoid metabolism; for example, overexpression of the *AtMYB12* gene of *Arabidopsis* enhanced the accumulation of flavonoid content under low temperatures in a light-dependent manner (Bhatia et al., 2018). During strawberry ripening, the *MYB10* gene in *Fragaria* × *ananassa* plays a general regulatory role in the flavonoid/phenylpropanoid pathway (Puche et al., 2014). In addition, *ATMYB12* (Bhatia et al., 2018), *ATMYB018* (Ballesteros et al., 2001), *ATMYB21* (Shin et al., 2002), *ATMYB075* and *ATMYB090* (Li et al., 2006) genes are involved in the light signaling pathway in *Arabidopsis*. R2R3-MYB TF family members in petunia (*Petunia hybrida*) are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning (Albert et al., 2011).

Chinese jujube (*Ziziphus jujuba* Mill.) is a traditional economic tree species in China. Jujube has flourished for a long time because of its strong resistance, simple management, high yield, rich nutrition, and good economic and ecological benefits (Li et al., 2007; Zhao, Liu & Tu, 2008). Many consumers favor jujube because of its good taste, rich nutrition, powerful health functions and alternative medicinal properties (Wang & Xu et al., 2018; Zhang & Bian et al., 2018; Lam et

al., 2016). Fresh jujube fruits develop rapidly particularly in the high-quality cultivar 'Dongzao' (Yuan et al., 2017). The entire Chinese jujube genome sequence was obtained in 2014 (Liu *et al.*, 2014), providing data for genome-wide analyses of the MYB superfamily. Although many studies have shown that MYB TFs are involved in physiological and biochemical processes in plants, the MYB TFs of jujube have not been fully identified and characterized, and the expression of *MYB* genes in the different developmental stages of fresh jujube fruit remains unclear. Thus, in this study, the MYB superfamily members in jujube were analyzed and identified, and their protein physicochemical properties, motif and exon/intron composition, correlation with flavonoid content, syntenic relationships, and expression levels in different developmental stages of fresh jujube were analyzed. These findings should inform the characterization of *ZjMYB* genes, and this study will be helpful for future functional studies of the *ZjMYB* superfamily genes involved in fruit development in jujube.

Materials and methods

Plant materials

Twelve-year-old Chinese jujube served as test materials at the 'Xuefeng' ecological park of Gengjiaying township, Yiliang county in Kunming city of Yunnan province, China, which is the teaching experiment base of Southwest Forestry University. Fruits were selected from five periods, including the young stage, enlargement stage, white mature stage, half-red stage, and full-red stage. The samples of these five periods were harvested in turn 25, 39, 74, 83 and 99 days after anthesis. These fruits were immediately frozen in liquid nitrogen and stored at -80°C. Three biological repeats were performed for each developmental stage for RNA-sequencing, total flavonoid content determination, and qRT-PCR analysis.

Jujube *MYB* superfamily gene identification

The completed jujube genome sequence and chromosome information were identified from the DDBJ/EMBL/Gen-Bank (accession JREP00000000), which provides Chinese jujube genome data (Liu *et al.*, 2014). The Pfam database (http://pfam.xfam.org/) was searched to obtain the hidden Markov model (HMM) profile for the MYB binding domain (PF00249) (Finn et al., 2016; Zhang & Ma et al., 2018), and all putative MYB genes were obtained from the jujube genome database (Liu *et al.*, 2014). The presence of a MYB domain in the selected MYB proteins was further verified using the online program SMART (http://smart.embl-heidelberg.de/) and HMMER (https://www.ebi.ac.uk/Tools/hmmer/). Next, manual analysis was performed using the ClustalX program to confirm MYB conserved domains or motifs. Ultimately, we confirmed that the genes containing the MYB domain were members of the MYB superfamily. Based on previous research, we downloaded 198 MYB family protein sequences and chromosome information in *Arabidopsis* from The *Arabidopsis* Information Resource (TAIR) (http://www.*Arabidopsis*.org/) (Stracke, Werber & Weisshaar, 2001; Chen et al., 2006).

Multiple sequence alignment and phylogenetic tree construction of MYB proteins

In this study, we will refer to the classification of the *Arabidopsis* MYB gene family and the results of a phylogenetic tree for subfamily classification of the jujube MYB family and grouping (Dubos et al., 2010; Chen et al., 2006; Zhang & Ma et al., 2018). We constructed the first phylogenetic tree with 126 R2R3-MYB proteins, 5 R1R2R3-MYB proteins and 1 4R-MYB protein of *Arabidopsis* and 171 MYB family proteins of jujube. We used the ClustalW program for multiple comparisons of amino acid sequences. The neighbor-joining (NJ) phylogenetic tree was constructed by 1,000 repeated bootstrap analyses using MEGA 6 software (Tamura et al., 2013). Genes not belonging to R2R3-MYB, R1R2R3-MYB and 4R-MYB in the first phylogenetic tree of MYB genes were used to construct a second phylogenetic tree with the 60 1R-MYB proteins and 6 atypical MYB proteins in *Arabidopsis*. Phylogenetic tree construction parameters reference

158 the first phylogenetic tree.

159

160 Gene structure and conserved motif analysis

161 The exon-intron structures of *ZjMYB* genes were generated with the Gene Structure Display Server
162 (GSDS: http://gsds.cbi.pku.edu.cn/) (Hu et al., 2015). The conserved motifs of *ZjMYB* proteins
163 were defined using the MEME program (http://meme-suite.org/tools/meme) (Bailey et al., 2009).
164 The following parameter settings were used: distribution of motifs, 0 or 1 per sequence; the
165 minimum width and maximum width of motifs, 6 and 250, respectively; maximum number of
166 motifs to find, 15; and default parameters. The conserved domains of *ZjMYB* proteins were defined
167 using Pfam (http://pfam.xfam.org/). We visualized the Pfam, MEME and GSDS results using
168 TBTOOLS software (http://cj-chen.github.io/TBtools/) (Chen et al., 2018; Liu et al., 2017). The
169 ExPASy online tool (http://www.expasy.ch/tools/protparam.html) was used to analyze the
170 physiological and biochemical characteristics of the *ZjMYB* genes.

171

172 Distribution on the chromosome and tandem duplication analysis of *MYB* genes
173 in jujube

174 The genome location of each *ZjMYB* member and the length of each chromosome of jujube was
175 obtained from the Chinese jujube database (Liu et al., 2014). Based on the above information, the
176 *ZjMYB* genes were mapped to corresponding locations on the twelve chromosomes using the circle
177 gene viewer package of TBtools (Krzywinski et al., 2009; Chen et al., 2018). Tandem duplication
178 analysis of *ZjMYB* genes was performed using the MCScanX tool (Wang et al., 2012).

179

180 Analysis of *ZjMYB* gene expression from RNA-Seq data

181 RNA-Seq reads were obtained with an Illumina HiSeq 2000. The fragments per kilobase of exon
182 per million mapped reads (FPKM) values were calculated based on RNA-Seq reads. The heatmap

was generated with TBTOOLS software (Chen et al., 2018; Liu et al., 2017); the color scale shown represents FPKM counts, and the ratios were log2 transformed. To confirm the transcriptome data, we chose 10 *MYB* genes and quantified them using quantitative real-time PCR experiments. Total RNA was extracted using the Plant RNA Kit (TaKaRa Biotechnology Co. Ltd., Dalian, China). DNA-free RNA was used for synthesis of the first strand of cDNA by using the Prime Script II 1st Strand cDNA Synthesis Kit (TaKaRa Biotechnology Co. Ltd., Dalian, China) per the manufacturer's recommendations. Quantitative RT-PCR was carried out with the Rotor-Gene Qreal-time PCR system (Qiagen, Germany) instrument using Fast Super EvaGreen qPCR Master Mix (US Everbright Inc., Suzhou, China). The gene-specific primers of the 10 ZjMYB genes were designed using Primer Premier 5.0 software (Lalitha, 2000), and the specific primer pairs are listed in Table S1. The UBQ gene of jujube was used as an internal control. Each reaction contained 10 μL of Fast Super EvaGreen qPCR Master Mix (US Everbright Inc., Suzhou, China), 0.4 μL of each primer, 2 μL of cDNA, and 7.2 μL of H_2O for a final volume of 20 μl according to the manufacturer's instructions. The reaction was carried out as follows: 95°Cfor 2min, followed by 40 cycles of 95°C for 5 s, 60°C for 5 s, 72°C for 25 s. Each reaction was performed using three-step amplification with three technical replicates, and the data from qRT-PCR amplification were analyzed using the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001).

Total flavonoid determination

The determined total flavonoid content was based on the colorimetric assay method at 510 nm, with slight modifications (Kou et al., 2015; Yu et al., 2012). Briefly, different concentrations (8, 16, 20, 32, 48, 64, and 80 mg/L) of a standard solution of rutin or extract was mixed with 0.3 mL of 5% $NaNO_2$. The solution was left to stand for 6 min, and then 0.3 mL of 10% $Al(NO_3)_3$ was added. After 6 min, 4 mL of 4% NaOH was added. Then, the mixture was made to 10 mL with distilled water. The absorbance of the final mixture was measured at 510 nm against a prepared blank using a spectrophotometer. The total flavonoid content was analyzed using OriginPro 8.5

software.

Correlation analysis

We performed a correlation analysis of the original FPKM counts to determine the expression levels of MYB superfamily genes with the total flavonoid content. Pearson correlation analysis was performed using the corr.test function of R software, and the coroplot software package of R was used for significance tests and graphing. $P<0.05$ was considered significant.

ZjMYB gene synteny analysis

Synteny analysis of *ZjMYB* genes between jujube and two plant species (peach and *Arabidopsis*) was performed using the MCScanX tool (Wang et al., 2012). A syntenic analysis map was constructed using TBtools software (Chen et al., 2018; Liu et al., 2017). The protein sequences and the genome location of each MYB gene and the length of each chromosome of *Arabidopsis* were downloaded from the TAIR database (http://www.arabidopsis.org/) (Swarbreck et al., 2008). Peach protein sequences and the genome location of each MYB gene and the length of each chromosome were downloaded from the phytozome database (https://phytozome.jgi.doe.gov/) (Zhang & Ma et al., 2018).

Results

Identification of jujube *MYB* genes and analysis of their protein physicochemical properties

A total of 171 *ZjMYB* genes with typical MYB or MYB-like domains were selected from the jujube genome database on the basis of the HMM profile of the MYB domain. The corresponding chromosomal locations were provisionally ordered by name such that the *MYB* genes were named *ZjMYB1* through *ZjMYB171*. The geneIDs, gene lengths, and physiological and biochemical

characteristics of the *ZjMYB* genes are listed in Table S2. The lengths of the coding sequences (CDSs) of the *ZjMYB* genes ranged from 276 to 3645 bp. The lengths of the protein sequences of *ZjMYB* genes ranged from 91 to 1214 amino acids. The isoelectric point (pI) values for *ZjMYB* proteins ranged from 4.78 (*ZjMYB108*) to 10.39 (*ZjMYB145*), with an average of 7.21, which is similar to those for MYB proteins in sesame (*Sesamum indicum* L.) and peach (Mmadi et al., 2017; Zhang & Ma et al., 2018). The molecular weight (Mw) ranged from 10558.01 Da (ZjMYB74) to 136172.41 Da (*ZjMYB23*), with an average of 42499.15 Da. The GRAVY（grand average of hydropathy）average value was -0.74, which is similar to that for peach MYB proteins (Zhang & Ma et al., 2018). The most likely homologous gene of each jujube MYB gene in *Arabidopsis* was obtained from the BLAST tool in the TAIR database (Table S2).

Phylogenetic trees and group classification of *ZjMYB* proteins

This study identified 171 *MYB* genes in jujube. As shown in Figures 1 and 2, we found 58 MYB-related proteins, 99 R2R3-MYB proteins, 4 R1R2R3-MYB proteins, 1 4R-MYB protein, and 9 atypical MYB proteins. R2R3-MYB proteins accounted for 58% of the *ZjMYB* proteins as the largest subfamily, and 4R-MYB subfamily proteins accounted for 0.6% as the smallest subfamily. MYB-related proteins accounted for 40% of *ZjMYB* proteins as the second largest subfamily.

The first phylogenetic tree (Fig. 1A) comprised 171 *ZjMYB* superfamily genes and 126 R2R3-MYB genes, 5 R1R2R3-MY*B* genes, and 1 4R-MYB gene of *Arabidopsis*. Fifty-eight MYB-related proteins and 9 atypical MYB proteins of the *ZjMYB* superfamily were clustered alone. Then, we used those genes with 60 MYB-related and 6 atypical MYB proteins of *Arabidopsis* to construct the second phylogenetic tree (Fig. 2A). As shown in Figure 1A, 1 jujube 4R-MYB protein (ZjMYB62) was clustered into the 4R-MYB protein of *Arabidopsis*, 4 jujube R1R2R3-MYB proteins were clustered into 5 R1R2R3-MYB proteins of *Arabidopsis*, and 99 jujube R2R3-MYB proteins were clustered into 126 R2R3-MYB proteins of *Arabidopsis*. These 99 R2R3-MYB genes of jujube were divided into 35 groups, C1 to C35, based on the topology of the tree and the classification of the MYB superfamily in *Arabidopsis*, peach and pear (*Pyrus bretschneideri*)

(Dubos et al., 2010; Zhang & Ma et al., 2018; Li X et al., 2016). The number of members in each group ranged from 1 to 7 in jujube. C10 had 7 members and was the largest group.

In the second phylogenetic tree (Fig. 2A), 58 MYB-related (1R-MYB) proteins were divided into 6 groups based on the topology of the tree and their classifications in *Arabidopsis* and peach (Chen et al., 2006; Zhang & Ma et al., 2018). CPC-like (3 members), TBP-like (11 members), I-box-binding-like (4 members), R-R-type (5 members), CCA1-like (13 members), and Chinese jujube-specific (22 members) groups were present.

TBP-like was the second largest group among the six groups of the MYB-related subfamily, as in peach (Zhang & Ma et al., 2018). The 22 members did not cluster with any *Arabidopsis* proteins. However, this result has been shown in previous studies in peach (Zhang & Ma et al., 2018) and sweet orange (Hou et al., 2014), and the authors included those genes in the MYB-related (1R-MYB) subfamily. The author of the peach study reported that those genes might have fruit-related functions and were either not needed in *Arabidopsis* or were obtained after divergence from the last common ancestor (Zhang & Ma et al., 2018). This study further contributes to this speculation as the function of those genes remains unknown. Nine atypical MYB proteins were clustered in 4 branches in the phylogenetic tree.

ZjMYB **gene structure and protein motif analysis**

The MEME program predicted 15 conserved motifs of the MYB proteins of jujube, which appeared in R2R3-MYB and R1R2R3-MYB (Table S3 and Fig. 1B), MYB-related and atypical MYB (Table S4 and Fig. 2B), and 4R-MYB (Table S5). The gene exon/intron structures were analyzed in R2R3-MYB with R1R2R3-MYB (Table S3 and Fig. 1C), MYB-related with atypical MYB (Table S4 and Fig. 2C), and 4R-MYB.

R2R3-MYB and R1R2R3-MYB had 3 identical highly conserved motifs (Fig. 1B), motif 3, motif 2, and motif 1. Seventy-eight of the 99 R2R3-MYB proteins had 4 identical highly conserved motifs, motif 3, motif 4, motif 2 and motif 1. As the sequence logos of the R2R3-MYB protein repeats show (Fig. 3), motif 3, motif 4, and the front part of motif 2 constitute the R2 repeat in

jujube and include a highly conserved separated triplet of Tryptophan (W) to maintain the HTH structure. This result was consistent with those of previous studies (Zhang & Ma et al., 2018; Hou et al., 2014; Chen et al., 2006) showing that the back parts of motif 2 and motif 1 constituted the R3 repeat. In this study, the first Tryptophan (W) of the R3 repeat in jujube was replaced by Leucinol hydrochloride (L), Isoleucine (I), Phenylalanine (F), Tyrosine (Y) and Methionine (M) (the back part of motif 2 in Fig. 3). The second and third Tryptophan residues were conserved in most R3 repeats in this study; only 1 (ZjMYB170) occurred in group C34 (2 members), and the third Tryptophan (W) residue was replaced by Phenylalanine (F) (the last amino acid of motif 1 in Fig. 3). Consistent with previous research, the R2 repeat also had the following highly conserved groups of EED residues (motif 3): glutamic acid (E)-glutamic acid (E)-aspartic acid (D) and EEE residues (motif 2) (Zhang & Ma et al., 2018; Li X et al., 2016). In jujube, the remaining residues were highly conserved in the R2 repeat; the eighth and sixteenth residues behind the first Tryptophan (W) were leucine (L) and glycine (G), and the eighth (G) and ninth (L) residues were behind the second Tryptophan (W). In front of the third Tryptophan (W) in the R2 repeats, the following 9 consecutive highly conserved residues were observed: arginine (R), cysteine (C), Glycine (G), Lysine (K), Serine (S), cysteine (C), arginine (R), leucine (L), and arginine (R). This characteristic was also found in other plant species, such as peach (Zhang & Ma et al., 2018) and Chinese white pear (Li X et al., 2016). Although the authors did not mention this characteristic in their articles, we found it by reanalyzing the data they presented. Other R2R3-MYB proteins (21 members) did not contain motif 4, and they were divided into 8 groups, including 1 in group C7 (3 members) and all members in group C29 to C35. Fifty-seven of the 99 R2R3-MYB proteins included motif 3, motif 4, motif 2 and motif 1 and also included motif 6. Although ZjMYB113 did not have motif 5 and ZjMYB42 did not have motif 6, groups C1, C2, C3, C9, and C7 had motifs 6, 3, 4, 2, 1 and 5. Groups C16 and C4 had motifs 6, 3, 4, 2, 1, 5 and 8. Motif 8 had highly conserved Tryptophan (W) residues, and the function of this W requires further research. The members of groups C11, C32, C30 and C34 had unique motifs, including 11, 12, 13 and 14 in jujube R2R3-MYB proteins.

316 Four R1R2R3-MYB proteins contained motifs 9, 3, 10, 2, 1 and 7. As the sequence logos show
317 (Fig. 4), motif 9 constituted the R1 repeat; motifs 3, 10 and 2 constituted the R2 repeat; and motifs
318 1 and 7 constituted the R3 repeat. The R2 repeat differed between R2R2-MYB and R1R2R3-MYB.
319 Motifs 3, 4 and 2 constituted the R2 repeat of R2R3-MYB.
320 The coding sequences of all R2R3-MYBs were disrupted by introns (Fig. 1C), except for group
321 32. Most coding sequences had three exons and two introns, and all of these exons included 2
322 short exons and 1 relatively long exon, similar to those in peach (Zhang & Ma et al., 2018). All
323 group 32 members had no introns and only 1 exon. However, these members had green regions
324 upstream and downstream of the front and back end of the sequences, which was consistent with
325 those in peach. Two members of group C34 had a maximum of 11 introns in the R2R3-MYB
326 subfamily. The members of groups 18 and 19 had 2 exons and 1 intron. The gene length, exon
327 number and exon/intron structure within the same group of R2R3-MYB were similar, which
328 supported the classification of R2R3-MYB subfamilies. R1R2R3-MYB introns ranged from 6 to
329 12, and ZjMYB134 had a maximal number of 12 introns.
330 Although comprehensive analyses of the motifs of R2R3-MYB proteins haves been performed in
331 previous studies, a similar comprehensive motif analysis for the MYB-related subfamily is lacking.
332 In this study, 15 motifs of 58 MYB-related proteins and 9 atypical MYB proteins are shown in
333 Figure 2B The groups contained a highly conserved motif 2 except the Chinese jujube-specific
334 group. The specific group consisted of at least conserved motifs 1 and 2, which together comprised
335 the R1 repeat. The members of the R-R-type group had motifs 2, 5, 2 and 3 in that order and
336 contained 2 repeats. Motifs 2 and 5 comprised the first repeat, and motifs 2 and 3 comprised the
337 second repeat. The CCA1-like group was divided into 2 parts consisting of at least conserved
338 motifs 2 and 3, which together comprised the R1 repeat. All CCA1-like members contained the
339 conserved motif SHAQK with motif 3 of the MYB repeat, consistent with previous studies in
340 *Arabidopsis* (Chen et al., 2006). Nine members of the second part of the CCA1-like group (12
341 members) also contained motif 8. Four members of the I-box-like group had motifs 2 and 5, which
342 together constituted the R1 repeat of the I-box-like group. The lengths of all I-box-binding-like

protein sequences were short and similar. The 11 members of TBP-like had motifs 2 and 7 except for *ZjMYB29*. Five members of 9 atypical MYB proteins had only motif 2, and other atypical MYB proteins had motifs 2 and 2 in that order, and they comprised an imperfect R1 repeat.

To clearly display the repeats of the 7 different groups in MYB-related proteins, we used the MEME program to define the motifs of 7 groups with highly conserved motifs (Fig. 5 and Table S6). The motifs in the sequence logos of R-R-type repeats had higher similarity than did those in *Arabidopsis* and peach (Chen et al., 2006; Zhang & Ma et al., 2018). The CCA1-like subfamily is composed of two main clades and is the same in *Arabidopsis* (Chen et al., 2006). The primary structure of clade Ⅰ repeat DNA-binding domains was [-W-(X 19)-W-], and the repeat of clade Ⅱ was [-W-(X 18)-W-]. The primary structure of the TBP-like repeat DNA-binding domains was [-W-(X 19)-W-]. Most R1 repeats of MYB-related proteins contained highly conserved groups of EED or EEE residues in jujube.

The coding sequences of the MYB-related subfamily were also disrupted by introns (Fig. 2C). The R-R-type and I-box-binding-like genes had 2 exons and 1 intron except for ZjMYB33 of the R-R type, which was similar to groups 18 and 19 of the R2R3-MYB subfamily. Most Chinese jujube-specific genes had 4 and 5 introns except for 3 genes that containing 6 introns and 1 that had 3 introns. On average, the CCA1-like group was disrupted by 4.8 introns, the CPC-like group had 13 introns, the atypical MYB group had 4.1 introns, and the TBP-like group had 4.8 introns.

We used the MEME program to define the motifs of 4R-MYB (*ZjMYB62*) and *Arabidopsis* (*AT3G18100.1*). Their motifs were very similar, and they all included highly conserved motifs 1 to 15. Seven groups had highly conserved motifs (Fig. 6A and Table S5), indicating that their sequences were similar. These groups contained 4 MYB DNA-binding domains. The jujube *4R-MYB* gene had 9 introns (Fig. 6B), and the first intron was the longest.

Distribution on the chromosome and tandem duplicated analysis of *MYB* genes

in jujube

Jujube has twelve chromosomes (Zj01 to Zj12) (Fig. 7), and 142 of 171 *ZjMYBs* unevenly distributed on chromosomes 1 to 12 of the jujube genome (Table S2 and Fig. 7). From Figure 7, we can find the higher density of the chromosome genes, and the number of *ZjMYBs* was also relatively large on each chromosome. Zj01 is the largest chromosome and had the most *ZjMYBs*. Three tandem duplicated pairs of jujube *MYB* genes were found on chromosome 2 (*ZjMYB25* with *ZjMYB26*), chromosome 5 (*ZjMYB64* with *ZjMYB65*) and chromosome 11 (*ZjMYB123* and *ZjMYB124*).

Analysis of *ZjMYB* gene expression from RNA-Seq data

The raw transcriptome sequences from RNA-Seq have been deposited into the NCBI sequence read archive (SRA) under accessions SRP162927. As shown in Table S7, a total of 46,464,880 - 50,619,884 raw reads were generated from the fifteen libraries. After filtration, a total of 46,244,076 - 50,550,982 clean reads were obtained from the fifteen libraries with an average Raw Q30 and Clean Q30 base rate of nearly 94%. The overall quality of the sequence data was suitable for further analysis. As revealed by jujube RNA-Seq (Fig. 8A and 8B), the expression levels of different genes at different developmental stages varied. TBtools software yielded a clearer picture of the trend of expression levels of different genes in different stages; the row was also log2 transformed. The reliability of the RNA-Seq data was further validated through real-time quantitative PCR (RT-qPCR) experiments which were carried out on 10 selected MYB genes at five stages (Fig. 8C).

Although extensive analysis of the expression levels of R2R3-MYB genes has been performed (Kranz, et al., 1998; Stracke, Werber & Weisshaar, 2001; Jia, Clegg & Jiang, 2004), a similar analysis of 1R-MYB (MYB-related) genes is limited. Interestingly, based on the original FPKM counts in the rectangles, the expression levels of most MYB-related genes (Fig. 8B) were relatively higher than those of R2R3-MYB, R1R2R3-MYB and 4R-MYB genes (Fig. 8A) in jujube fruit.

Based on the expression profiles of the *ZjMYB* family, the R2R3-MYB, R1R2R3-MYB and 4R-

MYB subfamilies can be divided into 9 subgroups, A1-A9 (Fig. 8A), and the MYB-related subfamily can be divided into 8 subgroups, B1 to B8 (Fig. 8B).

Interestingly, 4 genes (*ZjMYB126*, *ZjMYB7*, *ZjMYB10* and *ZjMYB72*) in group B5 and 2 genes (*ZjMYB92* and *ZjMYB127*) in group A7 similar changes in their gene expression levels, and all these genes showed a trend from high to low in Y to F. In contrast, the *ZjMYB34* gene in group A1 and 3 genes (*ZjMYB23*, *ZjMYB41* and *ZjMYB67*) in group B3 had similar gene expression changes from low to high in Y to FR. According to previous studies, the content of total flavonoids, phenolics and proanthocyanidins decrease as ripening progresses (Wu et al., 2012). Therefore, we predicted that the 10 genes of groups B3, B5, A1 and A7 may possess the function of activating flavonoids and proanthocyanidin biosynthesis.

In A7, most genes displayed high expression at stages Y and EN and had lower expression levels and even no expression in the other stages. Based on this finding, we can infer that the genes in group A7 probably participate in important regulatory functions at the early stages of jujube fruit development. In contrast, group A2 had relatively high expression levels at the late stages (WM, HR and FR). Compared with the genes in other groups, most genes in groups A1 (except for *ZjMYB34*), A3 (except for *ZjMYB99*), A4 (except for *ZjMYB88)*, A6, A8 (except for *ZjMYB136*), A9 (except for *ZjMYB48*), B4 (except for *ZjMYB63*) and B8 had lower expression levels, as well as no expression. The gene expression trend from Y to FR in group B3 and *ZjMYB34* in group A1 was from low to high. The expression level of group A2 increased during the early stages and decreased in the middle stage of jujube fruit development. These results indicated that *MYB* genes may participate in different regulatory mechanisms during jujube fruit development.

Total flavonoid determination

Consistent with previous studies, the total flavonoid content of jujube decreased with jujube fruit ripening (Wu et al., 2012) (Fig. 9). At the FR stage, the content (0.44 ± 0.03 mg RE/g FW) was similar to that in *Z. jujube* cv. Hupingzao (0.47 ± 0.06 mg RE/g FW) (Kou et al., 2015). The total flavonoid content was highest in the Y stage and reached 1.23 ± 0.03 mg RE/g FW.

Correlation analysis

To further explore the associations between the original FPKM counts of transcriptome data of 93 *ZjMYB* genes (containing 42 MYB-related, 7 atypical MYB, 40 R2R3-MYB, 3 R1R2R3-MYB and 1 4R-MYB genes) and total flavonoid contents, we performed a correlation analysis using R software (Fig. 10). Fifty-six MYB genes (23 MYB-related, 4 atypical MYB, 26 R2R3-MYB, 2 R1R2R3-MYB and 1 4R-MYB genes) presented significant (<0.05) correlations.

Among these genes, 7 MYB-related (*ZjMYB23*, *ZjMYB41*, *ZjMYB67*, *ZjMYB109*, *ZjMYB111*, *ZjMYB121* and *ZjMYB139*), 1 atypical MYB (*ZjMYB110*), eight R2R3-MYB (*ZjMYB4*, *ZjMYB9*, *ZjMYB34*, *ZjMYB44*, *ZjMYB45*, *ZjMYB107*, *ZjMYB135* and *ZjMYB142*) and one 4R-MYB gene (*ZjMYB62*) had a negative correlation with the total flavonoid content. Sixteen MYB-related genes (*ZjMYB2*, *ZjMYB7*, *ZjMYB10*, *ZjMYB11*, *ZjMYB18*, *ZjMYB63*, *ZjMYB72*, *ZjMYB76*, *ZjMYB95*, *ZjMYB98*, *ZjMYB100*, *ZjMYB117*, *ZjMYB119*, *ZjMYB126*, *ZjMYB132* and *ZjMYB133*), 3 atypical MYB genes (*ZjMYB84*, *ZjMYB122* and *ZjMYB125*), eighteen R2R3-MYB genes (*ZjMYB5*, *ZjMYB31*, *ZjMYB32*, *ZjMYB43*, *ZjMYB49*, *ZjMYB51*, *ZjMYB59*, *ZjMYB61*, *ZjMYB78*, *ZjMYB81*, *ZjMYB92*, *ZjMYB102*, *ZjMYB108*, *ZjMYB114*, *ZjMYB127*, *ZjMYB136*, *ZjMYB140* and *ZjMYB141*) and two R1R2R3-MYB (*ZjMYB79* and *ZjMYB134*) genes were positively correlated with total flavonoid contents.

ZjMYB gene synteny analysis

Arabidopsis is the most important model plant, and many *MYB* genes in *Arabidopsis* have been functionally well characterized. To further determine the phylogenetic mechanisms of the jujube MYB superfamily, we constructed a map based on a syntenic analysis of MYB genes between jujube with other species, including plants from the Rosaceae (peach) and Brassicaceae (*Arabidopsis*) as well as the paralogous *MYB* genes in jujube (Fig. 11) to predict the function of *ZjMYBs* and analyze the phylogenetic mechanisms of *MYB* genes. This process only highlighted

449 the *MYB* genes for each species. To find more homologous *MYB* gene pairs between jujube and
450 other species, their common gene pairs were all located using syntenic blocks possessing fewer
451 than 5 homologous gene pairs when we visualized the results using TBtools software.
452 The 5 pairs of paralogous *MYB* genes within jujube were thus composed of 9 jujube *MYB* genes
453 (Fig. 11 (red line) and Table S8). Fifty-seven pairs of orthologous *MYB* genes in jujube and
454 *Arabidopsis* were identified among the 171 *ZjMYBs* and 198 *AtMYBs* (Fig. 11 (green line) and
455 Table S8). Thirty-seven pairs of orthologous *MYB* genes in jujube and peach were identified
456 among the 171 *ZjMYBs* and 256 *PpMYBs* (Fig. 11 (blue line) and Table S8). Sixteen and 7 *MYB*
457 genes of jujube showed homology with multiple *MYB* genes from various locations on different
458 chromosomes in *Arabidopsis* and peach, respectively. Fourteen *ZjMYB* genes of jujube showed
459 the same homology to the *MYB* genes of *Arabidopsis* and peach.
460
461

462 # Discussion

463 A total of 171 *ZjMYB* genes were identified. Differences in the number of *MYB* genes among
464 diverse species have been studied. For instance, 256 *MYB* genes were found in peach (Zhang &
465 Ma et al., 2018), 177 in sweet orange (*Citrus sinensis*) (Hou et al., 2014) and 125 in physic nut
466 (Zhou et al., 2015). According to previous studies, the MYB family is the largest TF family in
467 jujube (Liu et al., 2017; Song et al., 2017; Shao et al., 2017; Zhang & Ma et al., 2017), consistent
468 with previous reports in *Arabidopsis* (Riechmann et al., 2000). The *ZjMYB* genes can be divided
469 into five subfamilies, including 58 MYB-related proteins, 99 R2R3-MYB proteins, 4 R1R2R3-
470 MYB proteins, 1 4R-MYB protein, and 9 atypical MYB proteins. The R2R3-MYB proteins make
471 up the largest subfamily, consistent with previous reports in peach and *Arabidopsis*. Similar to
472 jujube, a total of 128 R2R3-MYB proteins were identified in peach (Zhang & Ma et al., 2018), and
473 126 were found in *Arabidopsis* (Chen et al., 2006). Additionally, R2R3-MYB is the largest
474 subfamily of the MYB superfamily. However, this finding is inconsistent with the MYB-related
475 subfamily having the largest number of MYB superfamily genes in sesame (Mmadi et al., 2017).

These differences are most likely associated with the differences in the evolution of these plants. In this study, we identified 1 4R-MYB protein, which is consistent with previous studies of *Arabidopsis* (Dubos et al., 2010), peach (Zhang & Ma et al., 2018) and pear (Li X et al., 2016). However, the results of this study are inconsistent with those of previous studies reporting 4 4R-MYB proteins in Chinese cabbage (*Brassica rapa* ssp *pekinensis*) (Saha et al., 2016), 2 4R-MYB proteins in upland cotton (*Gossypium hirsutum* L.) (Salih et al., 2016), and 0 4R-MYB proteins in sesame (Mmadi et al., 2017). In this study, we identified 4 R1R2R3-MYB proteins, which is consistent with previous studies on peach (Zhang & Ma et al., 2018) and tomato (Li ZJ et al., 2016), but these results are inconsistent with other previous studies reporting 11 R1R2R3-MYB proteins in Chinese cabbage (Saha et al., 2016), 5 R1R2R3-MYB proteins in sesame (Mmadi et al., 2017) and 15 3R-MYB proteins in upland cotton (Salih et al., 2016). These differences are most likely associated with differences in the evolution of these plants.

In previous studies, the first Tryptophan (W) residue of the R3 repeat was replaced by Leucinol hydrochloride (L), Isoleucine (I) and Phenylalanine (F) (Ogata eta al., 1994; Du et al., 2012b). However, in this study, the first Tryptophan (W) of the R3 repeat in jujube was replaced by Tyrosine (Y) and Methionine (M) in addition to the above findings (the back part of motif 2 in Fig. 3); there were 2 (ZjMYB53 and ZjMYB54) in group C5 (3 members) and 1 (ZjMYB12) in group C12 (1 member). The second and third Tryptophan residues in this study were conserved in most R3 repeats, with 1 (ZjMYB170) occurring in group C34 (2 members), and the third Tryptophan (W) residue was replaced by Phenylalanine (F) (the last amino acid of motif 1 in Fig. 3). This finding might be attributed to the loss of residues during jujube evolution or a change in response to changes in the environment, further illustrating the diversity of the MYB domain (Du et al., 2012a). Consistent with previous research, the R2 repeat of R2R3-MYB genes also had highly conserved groups of EED and EEE residues (Zhang & Ma et al., 2018; Li X et al., 2016) (Fig. 3). In jujube, the other residues were highly conserved in the R2 repeat; the eighth and sixteenth residues behind the first Tryptophan (W) were Leucine (L) and Glycine (G), and the eighth (G) and ninth (L) residues were behind the second Tryptophan (W).

Although many comprehensive analyses of the motif for R2R3-MYB proteins have been performed in previous studies, a similar comprehensive analysis of the motif for the MYB-related subfamily is lacking. The 58 MYB-related (1R-MYB) proteins were divided into 6 groups, including the CPC-like, TBP-like, I-box-binding-like, R-R-type, CCA1-like, and Chinese jujube-specific groups (Fig. 2). Interestingly, the MYB repeats of CCA1-like proteins are closely related to R-R type MYB repeats (R-R(B)), and I-box-like protein MYB repeats are closely related to R-R-type MYB repeats (R-R(A)) (Fig. 5). We speculate that CCA1-like proteins and I-box-like proteins may be induced by R-R-type gene loss events.

The R2 repeat had highly conserved groups of EED and EEE residues in other plant species, such as peach (Zhang & Ma et al., 2018), pear (Li X et al., 2016) and tomato (Li ZJ et al., 2016). Interestingly, most R1 repeats of MYB-related proteins also included highly conserved groups of EED or EEE residues in jujube (Fig. 5).

The exon numbers of 171 ZjMYB genes ranged from 1 to 20 (Fig. 1C and Fig. 2C), indicating the loss and gain of ZjMYB exons during gene evolution, which may account for the functional diversity of the ZjMYB subgroup. As suggested by previous studies, the conserved motifs and intron/exon structures in each subgroup probably play important roles in each group's specific functions (Zhang & Ma et al., 2018; Li X et al., 2016).

From the Figure 7, we can see that more MYB genes are located at the two ends of the chromosome than on the middle of the chromosome, and similar patterns were also reported in peach (Zhang & Ma et al., 2018) and sweet orange (Hou et al., 2014). Despite the lack of description in the article, we can see from a previous paper that the MYB genes of *Arabidopsis* had similar distribution patterns on chromosome 1 and chromosome 5 (Katiyar et al., 2012).

The expression of most MYB-related genes (Fig. 8B) was relatively higher than that of R2R3-MYB, R1R2R3-MYB and 4R-MYB genes (Fig. 8A) in jujube fruit. Thus, we speculated that MYB-related genes probably play a greater role in fruit-related functions. Many of the functions of R2R3-MYB genes have been studied, but the function of MYB-related genes remains unknown. Therefore, we suggest further functional studies of MYB-related genes in the future.

Consistent with previous studies, the total flavonoid content of jujube decreased with jujube fruit ripening (Wu et al., 2012) (Fig. 9). According to previous studies, the MYB TF of *Arabidopsis* is involved in many secondary metabolic processes, such as the flavonoid metabolic pathway, glucosinolate biosynthesis, and anthocyanin biosynthesis (Stracke et al., 2010; Bhatia et al., 2018; Wei et al., 2015). For example, the *MdMYBPA1* TF of red-fleshed apple responded to low temperatures by redirecting the flavonoid biosynthetic pathway from proanthocyanidin to anthocyanin production (Liu et al., 2015). The MYB superfamily in other plants is also involved in flavonoid/phenylpropanoid metabolism (Puche et al., 2014; Ampomah -Dwamena et al., 2018; Wang & Qu et al., 2018). By correlation analysis between the original FPKM counts of transcriptome data from 93 *ZjMYB* genes (Fig. 10), 56 *MYB* genes presented a significant correlation.

Among the orthologous pairs (Fig. 11) (*ZjMYB108* with *AtMYB12* and *AtMYB11*; *ZjMYB46* with *AtMYB21*, *AtMYB24* and *AtMYB57*; *ZjMYB88* with *AtMYB15*; *ZjMYB69* with *AtMYB44*; *ZjMYB99* with *AtMYB44* and *AtMYB77*; *ZjMYB65* with *AtMYB28*; *ZjMYB107* with *AtMYB2*; *ZjMYB118* with *AtMYB62*;) in jujube and *Arabidopsis*, *AtMYB12* (*AT2G47460*) was expressed at low - temperature in a light-dependent manner, and overexpression of the *AtMYB12* gene enhanced flavonoid accumulation (Bhatia et al., 2018). *AtMYB11* and *AtMYB12* had a high degree of functional similarity and controlled flavanol biosynthesis (Stracke et al., 2010). We speculate that the *ZjMYB108* gene may have the same function. In the correlation analysis results (Fig. 10), *ZjMYB108* gene expression has a positive correlation with flavonoid content; therefore, we further speculate that *ZjMYB108* is an important transcription factor involved in the flavonoid metabolic pathway. *AtMYB21* (*AT3G27810*) is involved in the light signaling pathway (Shin et al., 2002). GA (gibberellin) promotes JA (jasmonate) biosynthesis to control the expression of *AtMYB21* (*AT3G27810*), *AtMYB24* (*AT5G40350*), and *AtMYB57* (*AT3G01530*) (Cheng et al., 2009). Transgenic *Arabidopsis* expressing *AtMYB15* improved tolerance to cold (Agarwal et al., 2006) and drought stress (Ding et al., 2008). *AtMYB44* (*AT5G67300*) was rapidly induced by MeJA (methyl jasmonate) in *Arabidopsis*. Transgenic *Arabidopsis* overexpressing the *AtMYB44* gene

was more sensitive to ABA and exhibited a markedly increased tolerance to salt and drought stress compared to wild-type plants (Jung et al., 2008). *MYB77* (*At3G50060*) regulates auxin signaling processes and auxin concentrations (Shin et al., 2007). *AtMYB28* (*AT5G61420*) is involved in aliphatic glucosinolate biosynthesis (Gigolashvili et al., 2007). Under drought stress in plants, the *AtMYB2* (*AT2G47190*) protein can function as a transcriptional activator in ABA-inducible gene expression (Abe et al., 2003). *AtMYB62* (*AT1G68320*) regulates phosphate starvation responses and gibberellic acid biosynthesis (Devaiah et al., 2009). The genes of the MYB orthologous pairs of jujube with *Arabidopsis* may have the same functions. We predict that the *ZjMYBs* are involved in flavonoid/phenylpropanoid metabolism, the light signaling pathway, auxin signal transduction, and responses to the various abiotic stresses (cold, drought and salt stresses). Fourteen *ZjMYB* genes of jujube had the same homology in the *MYB* genes of *Arabidopsis* and peach, indicating that the 14 *MYB* genes as well as the orthologous pairs with those 14 genes probably existed before the ancestral divergence of the MYB superfamily.

Conclusions

In this study, we performed the first genome-wide detailed analysis of jujube *MYB* superfamily genes, including 171 *MYB* genes (containing 58 MYB-related genes, 99 R2R3-MYB genes, 4 R1R2R3-MYB genes, 1 4R-MYB gene, and 9 atypical MYB genes). These 99 R2R3-MYB genes in jujube were divided into 35 groups, C1 to C35, and 58 MYB-related genes were divided into the following groups: the R-R-type, CCA1-like, I-box-binding-like, TBP-like, CPC-like and Chinese jujube-specific groups. The members of the *ZjMYB* gene superfamily in jujube were well supported by additional highly conserved motifs and exon/intron structures. Most R1 repeats of MYB-related proteins that also contained R2 repeats included highly conserved groups of EED and EEE residues in jujube. The composition of the motifs of the 4R-MYB protein in jujube was similar to that in *Arabidopsis*. The exon numbers of *ZjMYB* genes ranged from 1 to 20, indicating loss and gain of *ZjMYB* exons during gene evolution. Three tandem duplicated gene pairs were found on twelve chromosomes of jujube. The analysis of *ZjMYB* gene expression indicated that

584 the different genes had varying expression levels across developmental stages. Interestingly, in
585 this study, the expression of most MYB-related gene expression was higher than that of R2R3-
586 MYB, R1R2R3-MYB and 4R-MYB genes. In previous studies, the R2R3-MYB subfamily has
587 received more attention. However, in the present study, analysis of the expression of 126 *ZjMYB*
588 genes revealed that MYB-related genes played important roles in jujube development and
589 participated in fruit-related biological processes. The total flavonoid content of jujube decreased
590 with increasing ripening of jujube fruits. Ninety-three expressed genes were identified from the
591 RNA-sequencing data of jujube fruit, and 56 *ZjMYB* members presented a significant correlation
592 with the total flavonoid contents by correlation analysis. Five pairs of paralogous *MYB* genes
593 within jujube were thus composed of 9 jujube *MYB* genes. Fourteen *ZjMYB* genes of jujube showed
594 close homology to the *MYB* genes of *Arabidopsis* and peach, indicating that these 14 *MYB* genes
595 as well as their orthologs probably existed before the ancestral divergence of the MYB
596 superfamily. Based on a synteny analysis of *MYB* genes between jujube and *Arabidopsis*, we
597 predict that *ZjMYBs* are involved in flavonoid/phenylpropanoid metabolism, the light signaling
598 pathway, auxin signal transduction, and responses to various abiotic stresses (cold, drought and
599 salt stresses). We also speculate that *ZjMYB108* is an important transcription factor involved in the
600 flavonoid metabolic pathway. This study provided useful information that may serve as the basis
601 for functional analyses and cloning of *ZjMYB* genes. However, further studies are needed to
602 explore the functions of the *ZjMYB* genes to reveal the molecular regulatory of the mechanisms of
603 these genes in jujube fruit development.
604

605 References

606 Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchishinozaki K. 2003. *Arabidopsis AtMYC2* (bHLH)
607 and *AtMYB2* (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant Cell* 15(1),
608 63-78. DOI: 10.1105/tpc.006130
609 Agarwal M, Hao Y, Kapoor A, Dong CH, Fujii H, Zheng X, Zhu J. 2006. A R2R3 type MYB transcription
610 factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. *Journal of*

611 *Biological Chemistry* 281(49), 37636-37645. DOI: 10.1074/jbc.M605895200
612 Albert N W, Lewis D H, Zhang H, Schwinn KE, Jameson PE, Davies Km. 2011. Members of an
613 R2R3-MYB transcription factor family in *Petunia* are developmentally and environmentally regulated to
614 control complex floral and vegetative pigmentation patterning. *Plant Journal* 2011, 65(5):771-784. DOI:
615 10.1111/j.1365-313X.2010.04465.x
616 Allan AC, Hellens RP, Laing WA. 2008. MYB transcription factors that colour our fruit. *Cell* 13:99-102.
617 DOI: 10.1016/j.tplants.2007.11.012
618 Ampomah-Dwamena C, Thrimawithana AH, Dejnoprat S, Lewis D, Espley RV, Allan AC. 2018. A
619 kiwifruit (*Actinidia deliciosa*) R2R3-MYB transcription factor modulates chlorophyll and carotenoid
620 accumulation. *New Phytol*. Doi: 10.1111/nph.15362
621 Ballesteros ML, Bolle C, Lois LM, Moore JM, Vielle-Celzada JP, Grossniklaus U, Chua NH. 2001. LAF1,
622 a MYB transcription activator for phytochrome A signaling. *Genes & Development* 15(19):2613-25. DOI:
623 10.1101/gad.915001
624 Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME
625 SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37(Web Server issue):W202-W208.
626 DOI: 10.1093/nar/gkp33
627 Bhatia C, Pandey A, Reddy-Gaddam S, U Hoecker, Trivedi PK. 2018. Low Temperature Enhanced
628 Flavonol Synthesis Requires Light-Associated Regulatory Components in *Arabidopsis thaliana*. *Plant and*
629 *Cell Physiology*. DOI: 10.1093/pcp/pcy132
630 Chen CJ, Xia R, Chen H, He YH. 2018. TBtools, a Toolkit for Biologists integrating various HTS-data
631 handling tools with a user-friendly interface. *BioRxiv*. DOI: 10.1101/289660
632 Chen YH, Yang XY, He K, Liu MH, Li JG, Gao ZF, Lin ZQ, Zhang YF, Wang XX, Qiu XM, Shen YP, Li
633 Z, Deng XH, Luo JC, Deng XW, Chen ZL, Gu HY, Qu LJ. 2006. The MYB transcription factor superfamily
634 of *Arabidopsis thaliana*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant*
635 *Molecular Biology* 60:107-124. DOI:10.1007/s11103-005-2910-y
636 Cheng H, Song S, Xiao L, Soo H L, Cheng Z, Xie D, Peng J. 2009. Gibberellin Acts through Jasmonate to
637 Control the Expression of *MYB21*, *MYB24*, and *MYB57* to Promote Stamen Filament Growth in

Arabidopsis. *PLOS Genetics* 5(3). DOI: 10.1371/journal.pgen.1000440

Colanero S, Perata P, Gonzali S. 2018. The atroviolacea Gene Encodes an R3-MYB Protein Repressing Anthocyanin Synthesis in Tomato Plants. *Frontiers in Plant Science* 9:830. DOI: 10.3389/fpls.2018.00830

Cominelli E, Tonelli C. 2009. A new role for plant R2R3-MYB transcription factors in cell cycle regulation. *Cell Res* 19:1231-1232. DOI: 10.1038/cr.2009.123

Devaiah B N, Madhuvanthi R, Karthikeyan A S, Raghothama K G. 2009. Phosphate starvation responses and gibberellic acid biosynthesis are regulated by the *MYB62* transcription factor in *Arabidopsis*. *Molecular Plant* 2(1), 43-58. DOI: 10.1093/mp/ssn081

Ding Z, Li S, An X, Liu X, Qin H, Wang D. 2008. Transgenic expression of MYB15 confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*. *Cell Res* 18:1047-1060. DOI: 10.1016/S1673-8527(09)60003-5

Du H, Feng B R, Yang SS, Huang YB, Tang YX. 2012a. The R2R3-MYB transcription factor gene family in maize. *Plos One* 7(6):e37463. DOI: 10.1371/journal.pone.0037463

Du H, Yang SS, Liang Z, Feng BR, Liu L, Huang YB, Tang YX. 2012b. Genome-wide analysis of the MYB transcription factor superfamily in soybean. *Bmc Plant Biology* 12(1):106. DOI: 10.1186/1471-2229-12-106

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. 2010. MYB transcription factors in *Arabidopsis thaliana*. *Trends in Plant Science* 15:573-581. DOI: 10.1016/j.tplants.2010.06.005

Fang Q, Wang Q, Mao H, Xu J, Wang Y, Hu H, He S, Tu JC, Cheng C, Tian GZ, Wang XQ, Liu XP, Zhang C, Luo KM. 2018. AtDIV2, an R-R-type MYB transcription factor of *Arabidopsis*, negatively regulates salt stress by modulating ABA signaling. *Plant Cell Reports* 1-13. DOI: 10.1007/s00299-018-2321-6

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Batem an A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* 44:279-285. DOI: 10.1093/nar/gkv1344

Gigolashvili T, Yatusevich R, Berger B, Muller C, Flugge U. 2007. The R2R3-MYB transcription factor *HAG1/MYB28* is a regulator of methionine-derived glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant Journal* 51(2), 247-261. DOI: 10.1111/j.1365-313X.2007.03133.x

665 Hou XJ, Li SB, Liu SR, Hu CG, Zhang JZ. 2014. Genome-wide classification and evolutionary and
666 expression analyses of Citrus MYB transcription factor families in sweet orange. *PLos One* 9(11): e112375.
667 DOI: 10.1371/journal.pone.0112375
668 Hu B, Jin J, Guo AY, He Z, Luo J, Gao G. 2015. GSDS 2.0: an upgraded gene feature visualization server.
669 *Bioinformatics* 31(8):1296-7. DOI: 10.1093/bioinformatics/btu817
670 Jia L, Clegg MT, Jiang T. 2004. Evolutionary dynamics of the DNA-binding domains in putative R2R3-
671 MYB genes identified from rice subspecies indica and japonica genomes. *Plant Physiol* 134: 575-585. DOI:
672 10.1104/pp.103.027201
673 Jung C, Seo J S, Han SW, Koo Y J, Kim CH, Song S I, Nahm B H, Choi Y D, Cheong J J. 2008
674 Overexpression of *AtMYB44* enhances stomatal closure to confer abiotic stress tolerance in transgenic
675 *Arabidopsis*. *Plant Physiol* 146, 623-635. DOI: 10.1104/pp.107.110981
676 Katiyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V, Bansal KC. 2012. Genome-wide classification
677 and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genomics*
678 13:544. DOI: 10.1186/1471-2164-13-544
679 Klempnauer KH, Gonda TJ, Bishop JM. 1982. Nucleotide sequence of the retroviral leukemia gene v-myb
680 and its cellular progenitor c-MYB: the architecture of a transduced oncogene. *Cell 1982* 31:453-463. DOI:
681 10.1016/0092-8674(82)90138-6
682 Kou X, Chen Q, Li X, Li M, Kan C, Chen B, Zhang Y, Xue Z. 2015. Quantitative assessment of bioactive
683 compounds and the antioxidant activity of 15 jujube cultivars. *Food Chemistry* 173:1037-1044. DOI:
684 10.1016/j.foodchem.2014.10.110
685 Kranz HD, Denekamp M, Greco R, Jin H, Leyva A, Meissner RC, Petroni K, Urzainqui A, Bevan M,
686 Martin C, Smeekens S, Tonelli C, Paz-Ares J, Weisshaar B. 1998. Towards functional characterisation of
687 the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J* 16: 263-276. DOI:
688 10.1046/j.1365-313x.1998.00278.x
689 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones S, Marra M. 2009. Circos: an
690 information aesthetic for comparative genomics. *Genome Research* 19:1639-1645. DOI:
691 10.1101/gr.092759.109

692 Lalitha, S. 2000. Primer premier 5. *Biotechnol. Softw. Internet Rep* 1, 270-272.DOI:
693 10.1089/152791600459894
694 Lam CTW, Gong AGW, Lam KYC, Zhang LM, Chen JP, Dong TTX, Lin HQ, Tsim KWK. 2016. Jujube-
695 containing herbal decoctions induce neuronal differentiation and the expression of anti-oxidant enzymes in
696 cultured PC12 cells. *Journal of Ethnopharmacology* 188:275-283. DOI: 10.1016/j.jep.2016.05.015
697 Li J, Yang X, Wang Y, Li X, Gao Z, Pei M, Chen Z, Qu LJ, Gu H. 2006. Two groups of MYB transcription
698 factors share a motif which enhances trans-activation activity. *Biochemical & Biophysical Research*
699 *Communications* 341(4):1155-1163. DOI: 10.1016/j.bbrc.2006.01.077
700 Li JW, Fan LP, Ding SD, Ding XL. 2007. Nutritional composition of five cultivars of Chinese jujube. *Food*
701 *Chemistry* 103(2):454-460. DOI: 10.1016/j.foodchem.2006.08.016
702 Li KT, Zhang J, Kang YH, Chen MC, Song TT, Geng H, Tian J, Yao YC. 2018. McMYB10 Modulates the
703 Expression of a Ubiquitin Ligase, McCOP1 During Leaf Coloration in Crabapple. *Front Plant Sci* 9:704.
704 DOI: 10.3389/fpls.2018.00704
705 Li X, Xue C, Li J, Qiao X, Li L, Yu L, Huang Y, Wu J. 2016. Genome-wide identification, evolution and
706 functional divergence of MYB transcription factors in Chinese white pear (*Pyrus bretschneideri*). *Plant*
707 *and Cell Physiology* 57:824-847. DOI: 10.1093/pcp/pcw029
708 Li ZJ, Peng RH, Tian YS, Han HJ, Xu J, Yao QH. 2016. Genome-wide identification and analysis of the
709 MYB transcription factor superfamily in *Solanum lycopersicum*. *Plant and Cell Physiology* 57(8):1657-
710 1677. DOI: 10.1093/pcp/pcw091
711 Lipsick JS. 1996. One billion years of Myb. *Oncogene* 13:223-235. PMID:8710361
712 Liu C, Xie T, Chen C, Luan A, Long J, Li C, Ding Y, He Y. 2017. Genome-wide organization and
713 expression profiling of the R2R3-MYB transcription factor family in pineapple (*Ananas comosus*). *BMC*
714 *Genomics* 18(1):503. DOI: 10.1186/s12864-017-3896-y
715 Liu LS, White MJ, MacRae TH. 1999. Transcription factors and their genes in higher plants: functional
716 domains, evolution and regulation. *European Journal of Biochemistry* 262:247-257. PMID:10336605
717 Liu MJ, Zhao J, Cai QL, Liu GC, Wang JR, Zhao ZH, Liu P, Dai L, Yan GJ, Wang WJ, Li XS, Chen Y,
718 Sun YD, Liu ZG, Lin MJ, Xiao J, Chen YY, Li XF, Wu Bin, Ma Y, Jian JB, Yang W, Yuan Z, Sun XC,

719 Wei YL, Yu LL, Zhang C, Liao SG, He RJ, Guang XM, Wang Z, Zhang YY, Luo LH. 2014. The complex
720 jujube genome provides insights into fruit tree biology. *Nat Commun* 5:5315. DOI: 10.1038/ncomms6315
721 Liu Y, Shi Z, Maximova SN, Payne MJ, Guiltinan MJ. 2015. Tc-MYBPA an *Arabidopsis* TT2-like
722 transcription factor and functions in the regulation of proanthocyanidin synthesis in Theobroma cacao.
723 *BMC Plant Biology* 15(1):160. DOI: 10.1186/s12870-015-0529-y
724 Liu Z, Zhang L, Xue C, Fang H, Zhao J, Liu MJ. 2017. Genome-wide identification and analysis of MAPK
725 and MAPKK gene family in Chinese jujube (*Ziziphus jujuba* Mill.). *Bmc Genomics* 18(1):855. DOI:
726 10.1186/s12864-017-4259-4
727 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR
728 and the 2-ΔΔCT method. *Methods* 25: 402-408. DOI: 10.1006/meth.2001.1262
729 Mmadi MA, Dossa K, Wang LH, Zhou R, Wang YY, Cisse N, Sy MO, Zhang XR. 2017. Functional
730 characterization of the versatile MYB gene family uncovered their important roles in plant development
731 and responses to drought and waterlogging in sesame. *Genes* 8:362-380. DOI: 10.3390/genes8120362
732 Ogata K, Morikawa S, Nakamura H, Sekikawa A, Inoue T, Kanai H, Sarai A, Ishii S, Nishimura Y. 1994.
733 Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative
734 recognition helices. *Cell* 79(4):639-48. DOI: 10.1016/0092-8674(94)90549-5
735 Ogata K, Kaneiishii C, Sasaki M, Hatanaka H, Nagadoi A, Enari M, Nakamura H, Nishimura Y, Ishii S,
736 Sarai A. 1996. The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA
737 recognition and trans-activation. *Nature Structural Biology* 3:178-187. DOI: 10.1038/nsb0296-178
738 Paz-Ares J, Ghosal D, Wienand U, Peterson P, Saedler H. 1987. The regulatory c1 locus of Zea mays
739 encodes a protein with homology to MYB oncogene products and with structural similarities to
740 transcriptional activators. *EMBO Journal* 6:3553-3558. DOI: 10.1002/j.1460-2075.1987.tb02684.x
741 Puche LM, Cumplido-Laso G, Amil-Ruiz F, Hoffmann T, Ring L, Rodríguez-Franco A, Caballero JL,
742 Schwab W, Muñoz-Blanco J, Blanco-Portales R. 2014. MYB10 plays a major role in the regulation of
743 flavonoid/phenylpropanoid metabolismduring ripening of Fragaria × ananassa fruits. *Journal of*
744 *Experimental Botany* 65(2):401-417. DOI: 10.1093/jxb/ert377
745 Raffaele S, Vailleau F, Leger A, Joubes J, Miersch O, Huard C, Blee E, Mongrand S, Domergue F, Roby

D. 2008. A MYB transcription factor regulates very-long-chain fatty acid biosynthesis for activation of the hypersensitive cell death response in *Arabidopsis*. *The Plant Cell* 20(3), 752-767. DOI: 10.1105/tpc.107.054858

Riechmann JL, Heard J, Martin G, Reuber L, Jiang CZ, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G. 2000. *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290:2105-2110. PMID: 11118137

Saha G, Park JI, Ahmed NU, Kayum MA, Kang KK, Nou IS. 2016. Characterization and expression profiling of MYB transcription factors against stresses and during male organ development in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Plant Physiology and Biochemistry* 104:200-215. DOI: 10.1016/j.plaphy.2016.03.021

Salih H, Gong W, He S, Sun G, Sun J, Du X. 2016. Genome-wide characterization and expression analysis of MYB transcription factors in *Gossypium hirsutum*. *BMC Genet* 17:129. DOI: 10.1186/s12863-016-0436-8

Shao F, Qiang L, Wilson I W, Qiu D. 2017. Genome-wide identification and characterization of the SPL, gene family in *Ziziphus jujuba*. *Gene* 627:315-321. DOI: 10.1016/j.gene.2017.06.044

Shin B, Choi G, Yi H, Yang S, Cho I, Kim J, Lee S, Paek NC, Kim JH, Song PS, Choi G. 2002. AtMYB21, a gene encoding a flower-specific transcription factor, is regulated by COP1. *Plant Journal* 30(1):23-32. DOI: 10.1046/j.1365-313X.2002.01264.x

Shin R, Burch A Y, Huppert K, Tiwari, S B, Murphy, A S, Guilfoyle T J, Schachtman D P. 2007. The *Arabidopsis* Transcription Factor *MYB77* Modulates Auxin Signal Transduction. *The Plant Cell* 19(8), 2440-2453. DOI: 10.1105/tpc.107.050963

Singh K, Foley RC, Onate-Sanchez L. 2002. Transcription factors in plant defense and stress response. *Current Opinion in Plant Biology* 5(5):430-436. DOI: 10.1016/S1369-5266(02)00289-3

Song S, Zhou H, Sheng S, Cao M, Li Y, Pang X. 2017. Genome-Wide Organization and Expression Profiling of the SBP-Box Gene Family in Chinese Jujube (*Ziziphus jujuba* Mill.). *International Journal of Molecular Sciences* 18(8):1734. DOI: 10.3390/ijms18081734

773 Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, Weisshaar B. 2010. Differential
774 regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different
775 parts of the *Arabidopsis thaliana* seedling. *Plant Journal* 50(4):660-677. DOI: 10.1111/j.1365-
776 313X.2007.03078.x
777 Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current*
778 *Opinion in Plant Biology* 4:447-456. DOI: 10.1016/S1369-5266(00)00199-0
779 Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li DH, Meyer T,
780 Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P and Huala E. 2008. The
781 *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Res*
782 36:D1009-D1014. DOI: 10.1093/nar/gkm965
783 Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics
784 analysis version 6.0. *Molecular Biology & Evolution* 30, 2725e2729. DOI: 10.1093/molbev/mst197
785 Wang N, Qu C Jiang S, Chen Z, Xu H, Fang H, Su M, Zhang J, Wang Y, Liu W, Zhang Z, Lu N, Chen X.
786 2018. The proanthocyanidin-specific transcription factor MdMYBPA1 initiates anthocyanin synthesis
787 under low temperature conditions in red-fleshed apple. *Plant journal*. DOI: 10.1111/tpj.14013
788 Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H. 2012. MCScanX:
789 a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*
790 40(7):e49. DOI: 10.1093/nar/gkr1293
791 Wang YG, Xu Y, Ma XP, Liu XF, Yang MJ, Fan WG, Ren HW, Efehi N, Wang XL, Zhu XQ. 2018.
792 Extraction, purification, characterization and antioxidant activities of polysaccharides from *Zizyphus*
793 *jujuba* cv. Linzexiaozao. *International Journal of Biological Macromolecules* 118(Pt B):2138-2148. DOI:
794 10.1016/j.ijbiomac.2018.07.059
795 Wei X, Liu K, Zhang Y, Feng Q, Wang L, Zhao Y, Li D, Zhao Q, Zhu X, Zhu X, Li W, Fan D, Gao Y, Lu
796 Y, Zhang X, Tang X, Zhou C, Zhu C, Liu L, Zhong R, Tian Q, Wen Z, Weng Q, Han B, Huang X, Zhang
797 X. 2015. Genetic discovery for oil production and quality in sesame. *Nature Communication* 6:8609-8622.
798 DOI: 10.1038/ncomms9609
799 Weston K. 1998. Myb proteins in life, death and differentiation. *Curr Opin Genet Dev* 8: 76-81. DOI:

10.1016/S0959-437X(98)80065-8

Wu CS, Gao QH, Guo XD, Yu JG, Wang M. 2012. Effect of ripening stage on physicochemical properties and antioxidant profiles of a promising table fruit 'pear-jujube' (*Zizyphus jujuba* Mill.). *Scientia Horticulturae* 148(148):177-184. DOI: 10.1016/j.scienta.2012.09.026

Xing C, Liu Y, Zhao L, Zhang S, Huang X. 2018. A novel MYB Transcription Factor regulates AsA synthesis and effects cold tolerance. *Plant Cell & Environmen.* DOI: 10.1111/pce.13387

Yu Y, Zhang S, Ren Y, Li H, Zhang X, Di J. 2012. Jujube preservation using chitosan film with nano-silicon dioxide. *Journal of Food Engineering* 113(3):408-414. DOI: 10.1016/j.jfoodeng.2012.06.021

Yuan CG, Huo C, Gui B, Liu JF, Chen YS. 2017. Facile phyto-mediated synthesis of silver nanoparticles using chinese winter jujube (*Ziziphus jujuba* Mill. cv. Dongzao) extract and their antibacterial/catalytic properties. *Iet Nanobiotechnology* 11(8):973-980. DOI: 10.1049/iet-nbt.2016.0242

Zhang C, Bian Y, Hou S, Li X. 2018. Sugar transport played a more important role than sugar biosynthesis in fruit sugar accumulation during Chinese jujube domestication. *Planta*. DOI: 10.1007/s00425-018-2971-1

Zhang C, Ma R, Xu J, Yan J, Guo L, Song J, Feng R, Yu M. 2018. Genome-wide identification and classification of MYB superfamily genes in peach. *Plos One* 13(6):e0199192. DOI: 10.1371/journal.pone.0199192

Zhang L, Zhao J, Feng C, Liu M, Wang J, Hu Y. 2017. Genome-wide identification, characterization of the MADS-box gene family in Chinese jujube and their involvement in flower development. *Scientific Reports* 7(1):1025. DOI: 10.1038/s41598-017-01159-8

Zhao ZH, Liu MJ, Tu PF. 2008 Characterization of water soluble polysaccharides from organs of Chinese Jujube (*Ziziphus jujuba* Mill. cv. Dongzao). *European Food Research & Technology* 226(5):985-989. DOI: 10.1007/s00217-007-0620-1

Zhou C, Chen Y, Wu Z, Lu W, Han J, Wu P, Chen Y, Li M, Jiang H, Wu G. 2015. Genome-wide analysis of the MYB gene family in physic nut (*Jatropha curcas* L.). *Gene* 572(1):63-71. DOI: 10.1016/j.gene.2015.06.072.

Figure 1

Phylogenetic tree of the R2R3-MYB, R1R2R3-MYB and 4R-MYB subfamilies of jujube and *Arabidopsis* and conserved motif and gene structure analysis of the R2R3-MYB and R1R2R3-MYB subfamily proteins of jujube.

(A) Phylogenetic tree of the MYB proteins of jujube and *Arabidopsis.* The sequences of the 171 MYB superfamily proteins of jujube and 132 *Arabidopsis* R2R3-MYB, R1R2R3-MYB and 4R-MYB proteins were aligned by ClustalW, and the phylogenetic tree was constructed using MEGA 6. The small white triangles represent the 171 jujube MYB proteins, and the small black triangles represent the 132 *Arabidopsis* MYB proteins. The names of each group are marked by English letters with Arabic numbers. (B) Distributions of conserved motifs in ZjMYB genes. The motifs of numbers 1-15 are indicated in different colored boxes. The sequence information of the motifs is provided in Table S2. (C) The exon-intron structure of jujube MYB genes. The yellow boxes and black lines indicate exons and introns, respectively. The green boxes at the two ends of the sequences indicate upstream and downstream regions, respectively. The MYB domains are highlighted by red boxes.

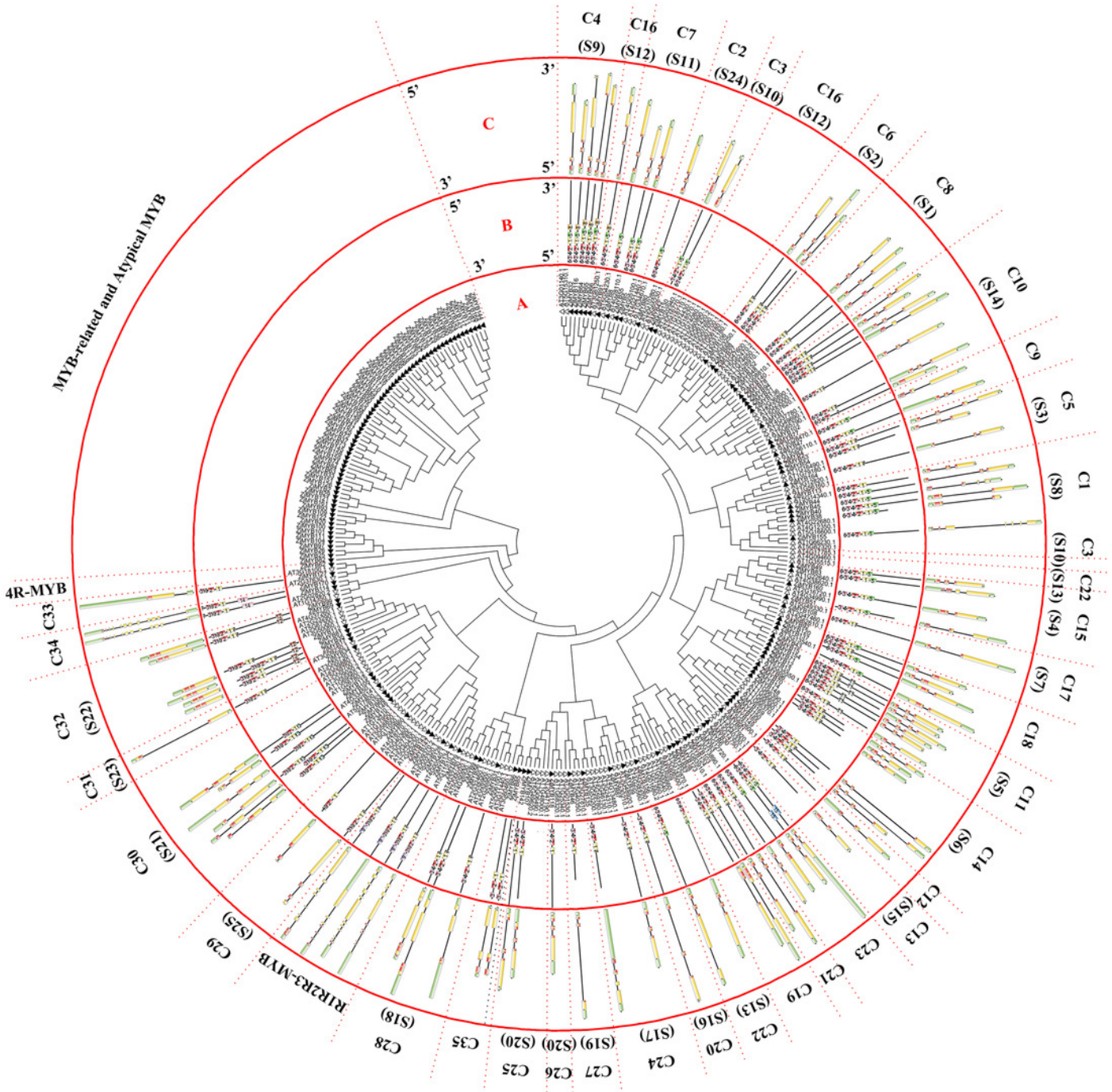

A
B
C
3'
5'
MYB-related and Atypical MYB
4R-MYB
RIKER3-MYB
C4 (S9)
C16 (S12)
C7 (S11)
C2 (S24)
C3 (S10)
C16 (S12)
C6 (S2)
C8 (S1)
C10 (S14)
C9
C5 (S3)
C1 (S8)
C3 (S10)(S13)
C22 (S4)
C15
C17 (S7)
C18
C11 (S5)
C14 (S6)
C12 (S15)
C13
C23
C21
C19
C22 (S13)
C20 (S16)
C24 (S17)
C27 (S19)
C26 (S20)
C25 (S20)
C35
C28 (S18)
C29 (S25)
C30 (S21)
C31 (S23)
C32 (S22)
C34
C33

Figure 2

Phylogenetic tree of the MYB-related (1R-MYB) and atypical MYB subfamilies of jujube and *Arabidopsis*, and conserved motif and gene structure analysis of MYB-related and atypical MYB subfamily proteins of jujube.

(A) Phylogenetic tree of the MYB proteins of jujube and *Arabidopsis*. The sequences of the 58 MYB-related and 9 atypical MYB proteins of jujube with the 60 MYB-related and 6 atypical MYB proteins of *Arabidopsis* were aligned by ClustalW, and the phylogenetic tree was constructed using MEGA 4. The small white triangles represent the 67 jujube MYB proteins and the small black triangles represent the 66 Arabidopsis MYB proteins. The names of each group are marked by English letters with Arabic numbers. (B) Distributions of conserved motifs in ZjMYB genes. The motifs of numbers 1-15 are indicated in different colored boxes. The sequence information of the motifs is provided in Table S3. (C) The exon-intron structure of jujube MYB genes. The yellow boxes and black lines indicate exons and introns, respectively. The green boxes at the two ends of the sequences indicate upstream and downstream regions, respectively. The MYB domains are highlighted by red boxes.

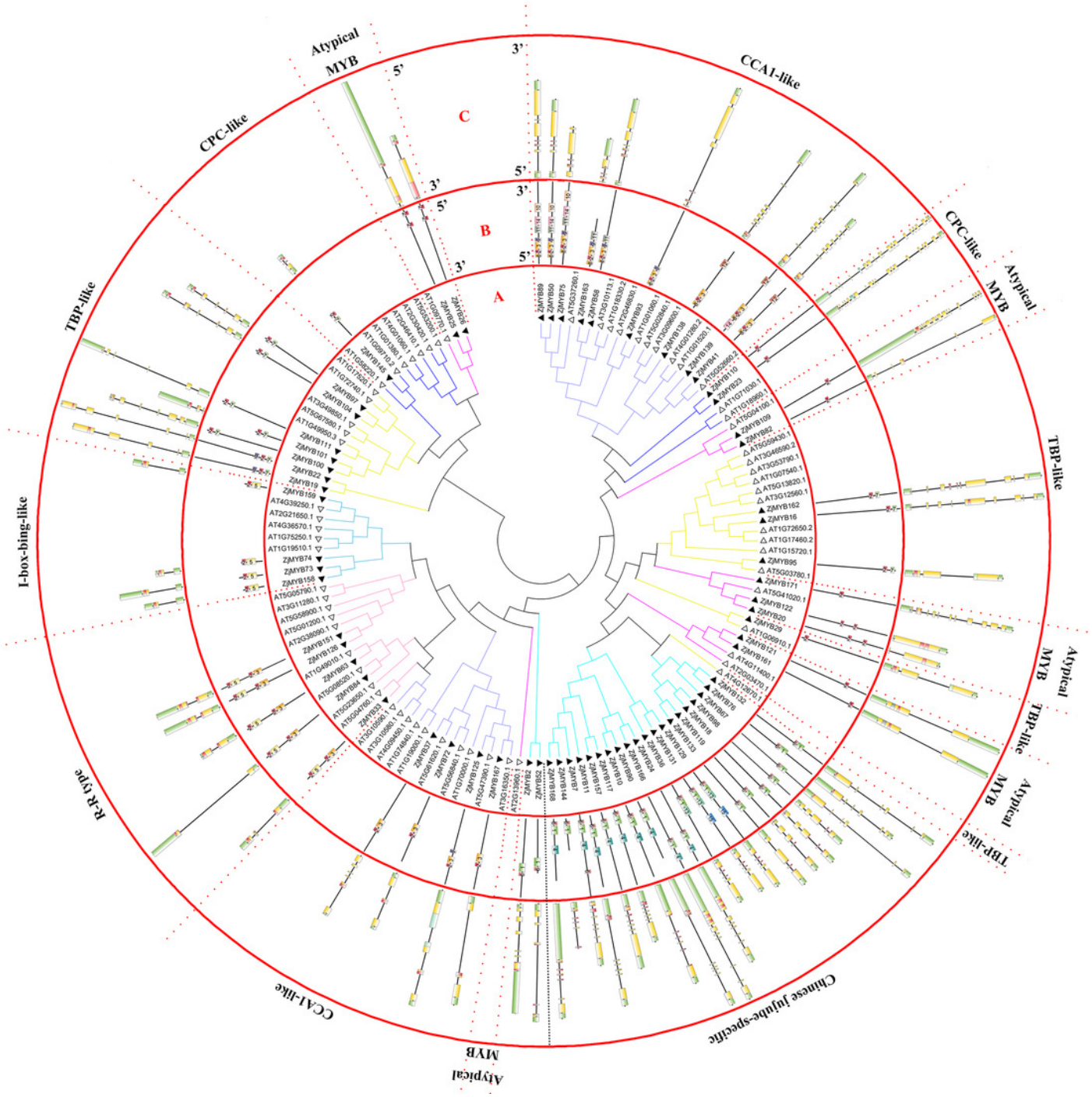

Atypical MYB
CCA1-like
CPC-like
Atypical MYB
TBP-like
Atypical MYB
TBP-like
Atypical MYB
Chinese jujube-specific
Atypical MYB
CCA1-like
R-R type
I-box-bing-like
TBP-like
CPC-like
A
B
C
5'
3'

Figure 3

R2 and R3 MYB repeats of the proteins of R2R3-MYB subfamily in jujube.

The logo sequences of R2 and R3 repeats were composed of motifs 3, 4, 2 and 1 in jujube. The overall height of each stack showed the conservation of the MYB protein sequence at that position. English letters indicate the different type of amino acid residue.

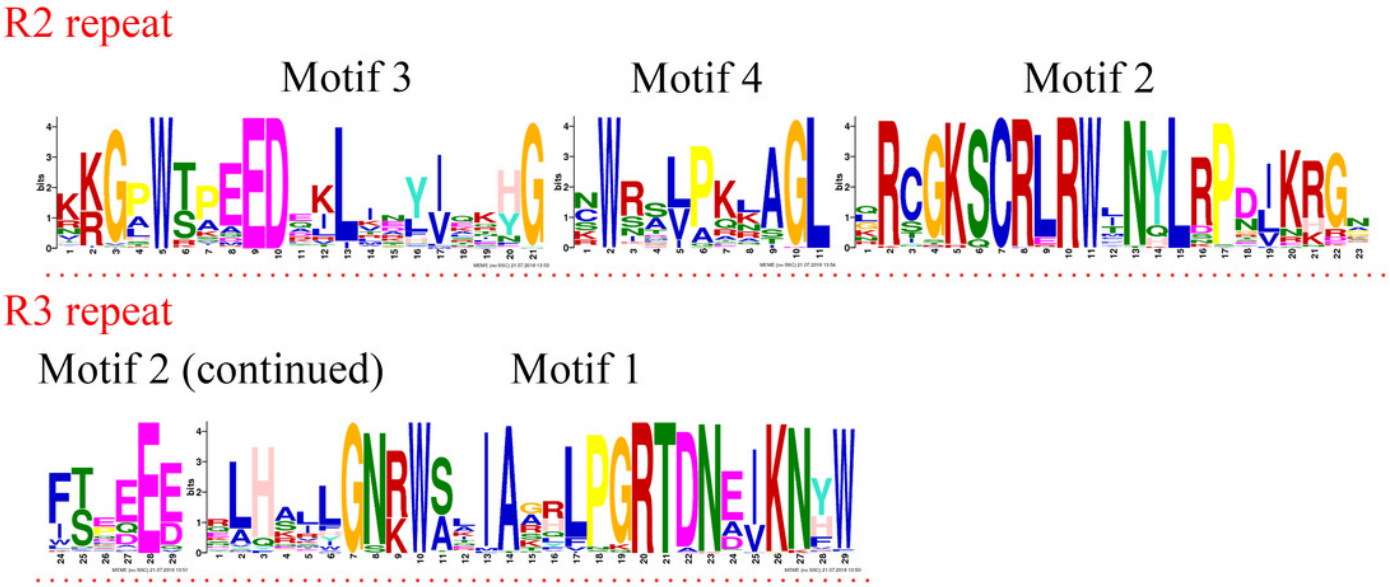


Figure 4

R1, R2 and R3 MYB repeats of the proteins of R1R2R3-MYB subfamily in jujube.

The logo sequences of R1, R2 and R3 repeats were composed of motifs 9, 3, 10, 2, 1 and 7 in jujube. The overall height of each stack showed the conservation of the MYB protein sequence at that position. English letters indicate the different type of amino acid residue.

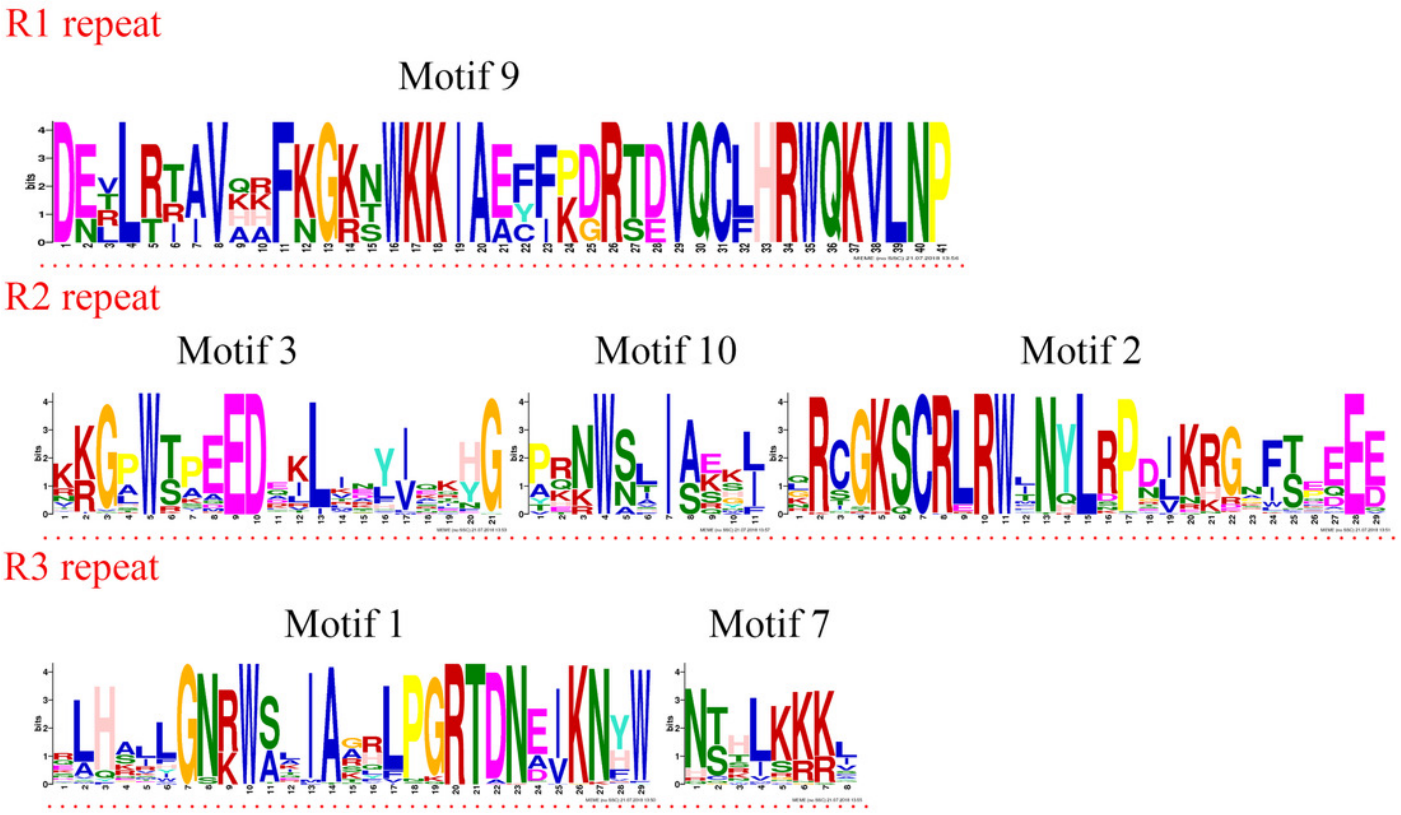


Figure 5

This logo indicates the sequence similarities of the ZjMYB repeats of the R-R-type, CCA 1-like and TBP-like MYB-related proteins.

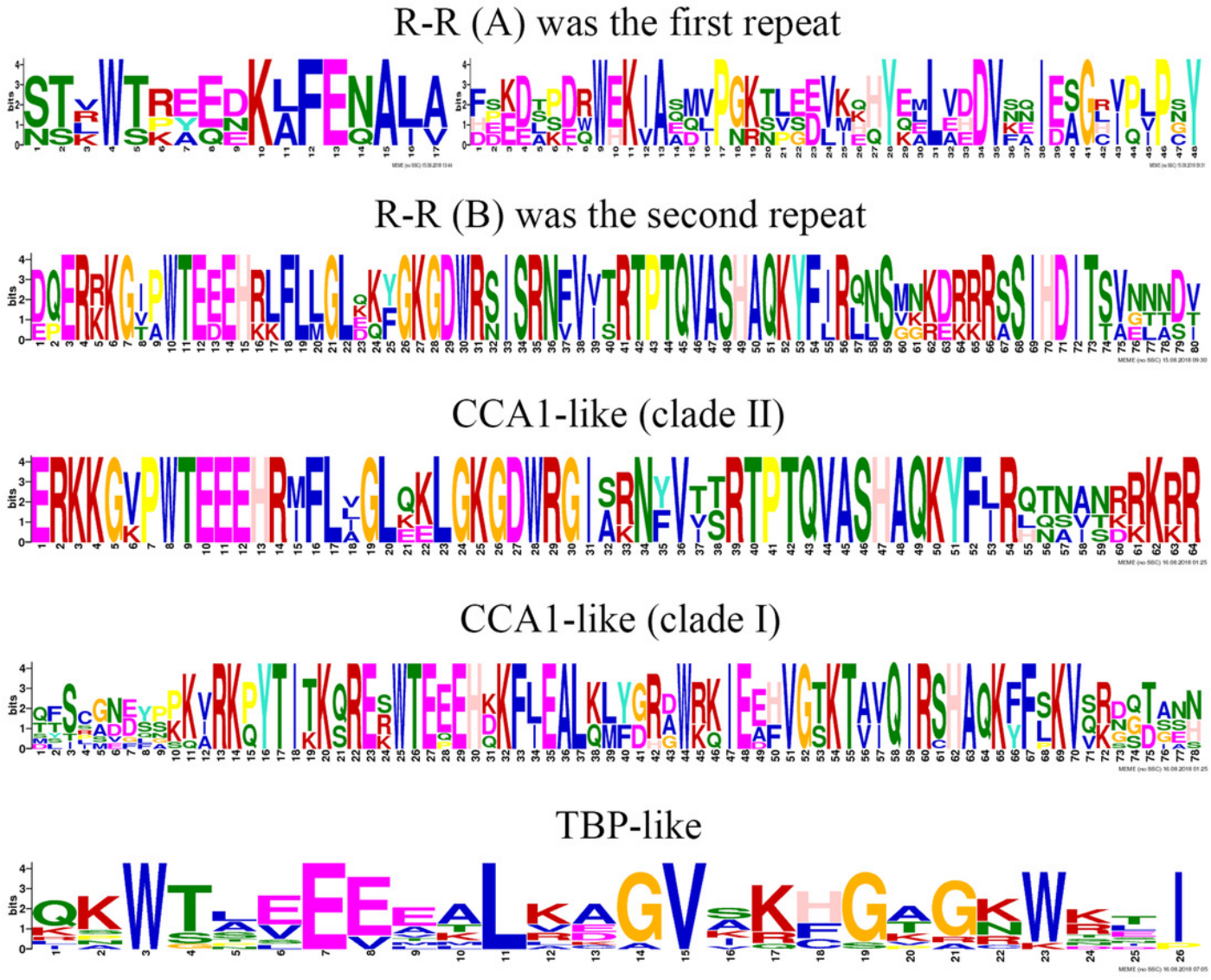


Figure 6

Motif and gene structure analysis of the 4R-MYB protein.

(A) Distributions of conserved motifs in 4R-MYB genes of *Arabidopsis* and jujube. The motifs of numbers 1-15, are indicated in different colored boxes. The sequence information of motifs had provided in Table S4. (B) Exon-intron structure of 4R-MYB in jujube. The yellow boxes and black lines indicate exons and introns, respectively. Green boxes at the two ends of sequences indicate upstream and downstream regions. The MYB domains are highlighted by red boxes.

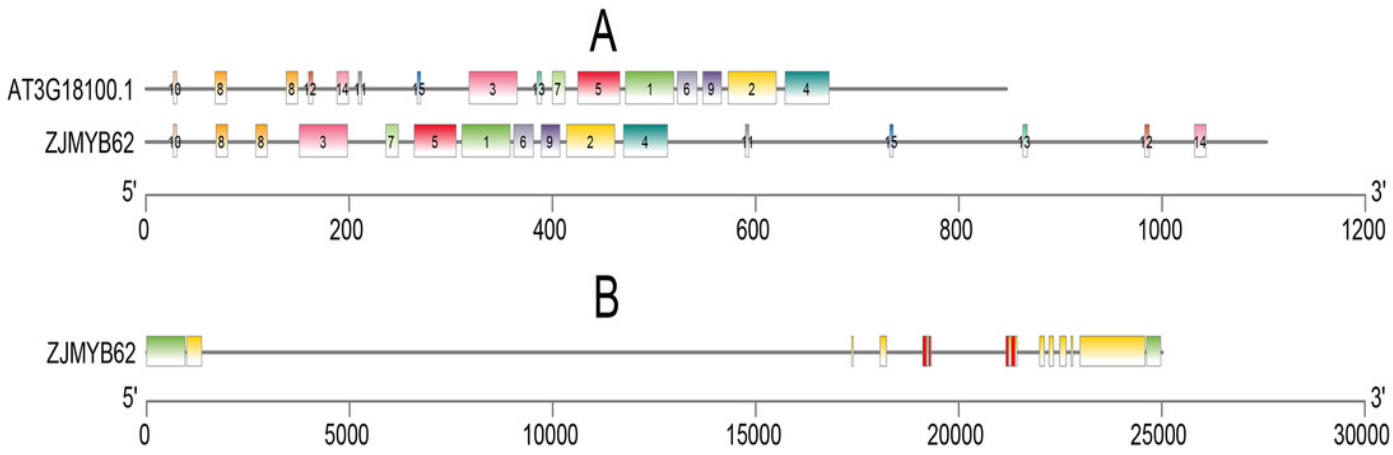


Figure 7

Chromosomal distribution and gene duplication of jujube MYB superfamily genes.

The inner ring denotes the 12 pseudochromosomes of jujube. The positions of the labels in the optical map are shown by the red bar (a higher bar corresponds to greater density). The outer ring shows the chromosomal distribution of jujube MYB superfamily genes. The scale is 5 Mb. The three yellow lines inside the circle are the tandem duplicated MYB gene pairs.

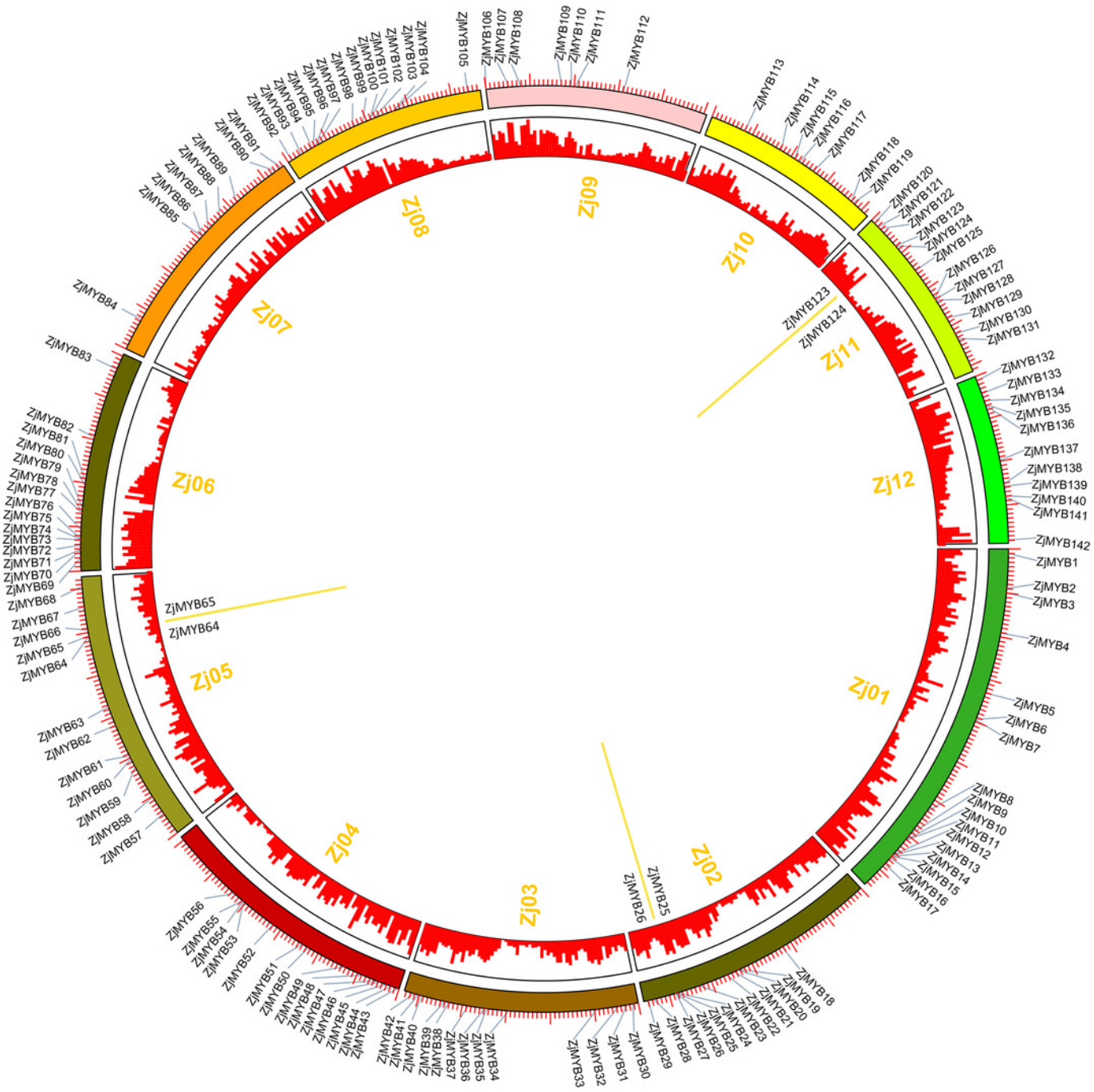

Zj01
Zj02
Zj03
Zj04
Zj05
Zj06
Zj07
Zj08
Zj09
Zj10
Zj11
Zj12
ZjMYB1
ZjMYB2
ZjMYB3
ZjMYB4
ZjMYB5
ZjMYB6
ZjMYB7
ZjMYB8
ZjMYB9
ZjMYB10
ZjMYB11
ZjMYB12
ZjMYB13
ZjMYB14
ZjMYB15
ZjMYB16
ZjMYB17
ZjMYB18
ZjMYB19
ZjMYB20
ZjMYB21
ZjMYB22
ZjMYB23
ZjMYB24
ZjMYB25
ZjMYB26
ZjMYB27
ZjMYB28
ZjMYB29
ZjMYB30
ZjMYB31
ZjMYB32
ZjMYB33
ZjMYB34
ZjMYB35
ZjMYB36
ZjMYB37
ZjMYB38
ZjMYB39
ZjMYB40
ZjMYB41
ZjMYB42
ZjMYB43
ZjMYB44
ZjMYB45
ZjMYB46
ZjMYB47
ZjMYB48
ZjMYB49
ZjMYB50
ZjMYB51
ZjMYB52
ZjMYB53
ZjMYB54
ZjMYB55
ZjMYB56
ZjMYB57
ZjMYB58
ZjMYB59
ZjMYB60
ZjMYB61
ZjMYB62
ZjMYB63
ZjMYB64
ZjMYB65
ZjMYB66
ZjMYB67
ZjMYB68
ZjMYB69
ZjMYB70
ZjMYB71
ZjMYB72
ZjMYB73
ZjMYB74
ZjMYB75
ZjMYB76
ZjMYB77
ZjMYB78
ZjMYB79
ZjMYB80
ZjMYB81
ZjMYB82
ZjMYB83
ZjMYB84
ZjMYB85
ZjMYB86
ZjMYB87
ZjMYB88
ZjMYB89
ZjMYB90
ZjMYB91
ZjMYB92
ZjMYB93
ZjMYB94
ZjMYB95
ZjMYB96
ZjMYB97
ZjMYB98
ZjMYB99
ZjMYB100
ZjMYB101
ZjMYB102
ZjMYB103
ZjMYB104
ZjMYB105
ZjMYB106
ZjMYB107
ZjMYB108
ZjMYB109
ZjMYB110
ZjMYB111
ZjMYB112
ZjMYB113
ZjMYB114
ZjMYB115
ZjMYB116
ZjMYB117
ZjMYB118
ZjMYB119
ZjMYB120
ZjMYB121
ZjMYB122
ZjMYB123
ZjMYB124
ZjMYB125
ZjMYB126
ZjMYB127
ZjMYB128
ZjMYB129
ZjMYB130
ZjMYB131
ZjMYB132
ZjMYB133
ZjMYB134
ZjMYB135
ZjMYB136
ZjMYB137
ZjMYB138
ZjMYB139
ZjMYB140
ZjMYB141
ZjMYB142

Figure 8

Heatmap of the expression levels of MYB genes in different developmental stages in jujube fruit and the expression levels of 10 ZjMYB genes by qRT-PCR.

Y (young stage), EN (enlargement stage), WM (white mature stage), HR (half-red stage) and FR (full-red stage). The color scale shown at the top represents log2-transformed FPKM (fragments per kilobase of exon per million mapped reads) counts. Original FPKM counts are displayed in the corresponding rectangles. Red indicates high expression and green indicates low expression. We eliminated the jujube MYB genes whose expression levels were 0 in all stages. (A) Heatmap of the expression levels of 71 R2R3-MYB, 3 R1R2R3-MYB and 1 4R-MYB genes in different developmental stages in jujube fruit. (B) Heatmap of the expression levels of 44 MYB-related and 7 atypical MYB genes in different developmental stages in jujube fruit. A1-A9 and B1-B8 indicate the different subgroups. (C) Expression analysis of 10 MYB genes in jujube by qRT-PCR. The data were normalized to the UBQ gene, and the vertical bars indicate the standard deviation.

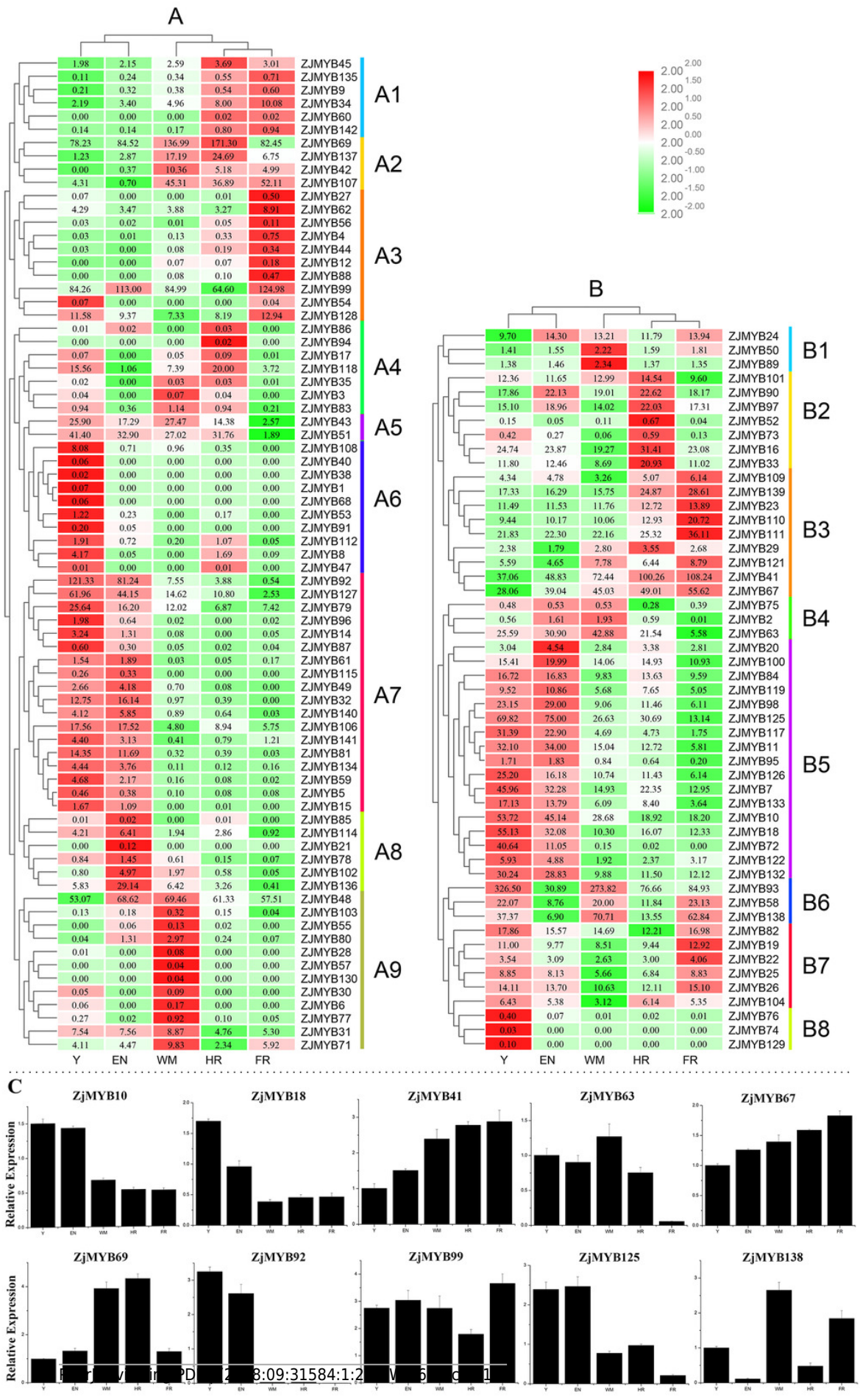

A
B
C
A1
A2
A3
A4
A5
A6
A7
A8
A9
B1
B2
B3
B4
B5
B6
B7
B8
Y
EN
WM
HR
FR
ZjMYB10
ZjMYB18
ZjMYB41
ZjMYB63
ZjMYB67
ZjMYB69
ZjMYB92
ZjMYB99
ZjMYB125
ZjMYB138
Relative Expression

Figure 9

Total flavonoids content of fruit development stages of jujube.

The results were represented as the mean ± standard deviation. Y (young stage), EN (enlargement stage), WM (white mature stage), HR (half-red stage) and FR (full-red stage).

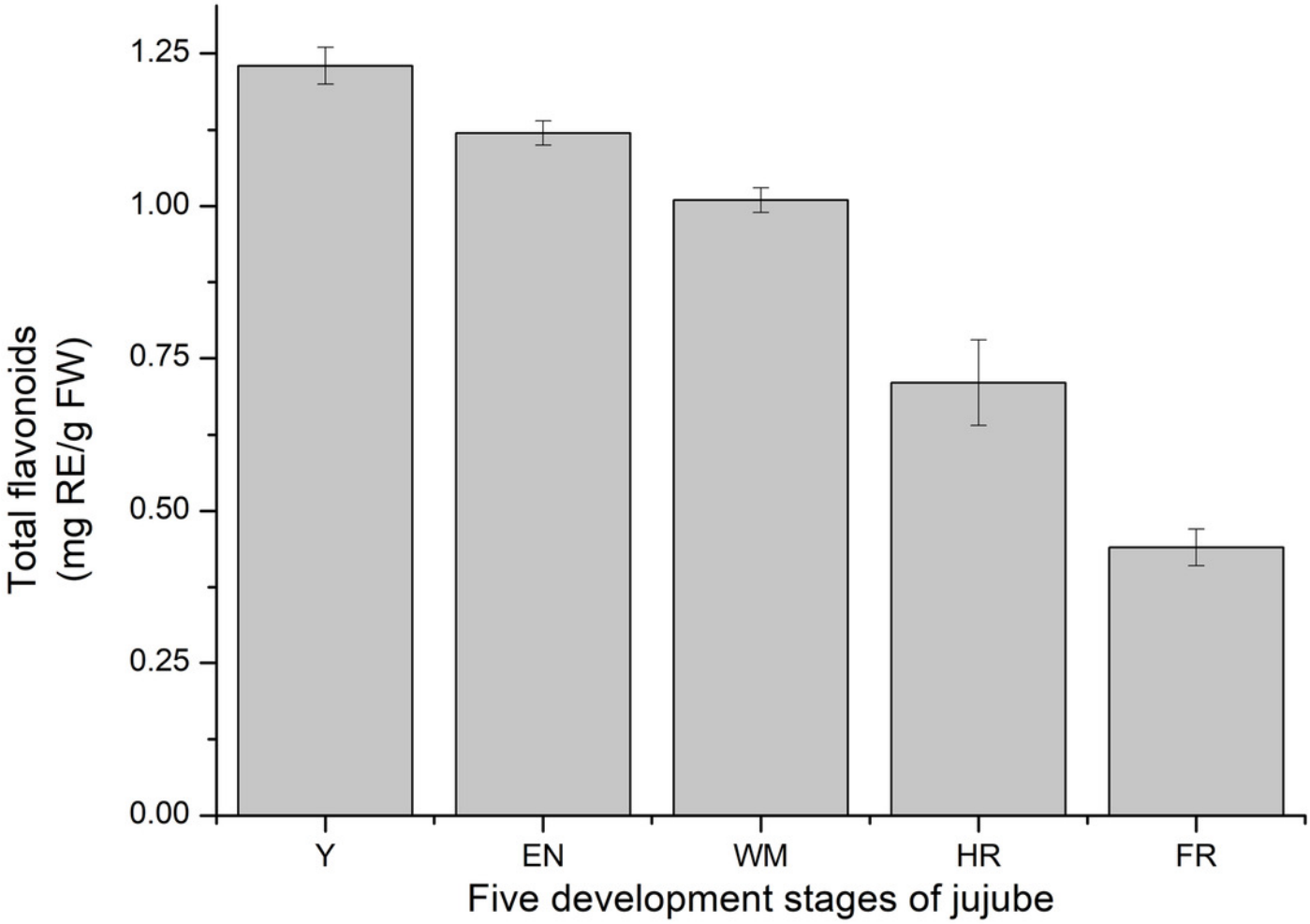


Figure 10

The correlation analysis of total flavonoid content with ZjMYB gene expression in jujube fruit.

The correlation analysis was performed using the original FPKM counts and the expression levels of the 93 MYB genes and the total flavonoid contents during the 5 developmental stages in jujube fruit. The "×" symbol indicates that the correlation is not significant at the 0.05 level. Yellow and red circles indicate that the correlation is significant and positive at the 0.05 level, and the green and blue circles indicate that the correlation is significant and negative at the 0.05 level. The correlations between total flavonoid contents and MYB genes are marked in the 2 red boxes. (A) The correlation analysis of total flavonoid contents with the 42 MYB-related and 7 atypical MYB genes. (B) The correlation analysis of total flavonoid contents with the 40 R2R3-MYB, 3 R1R2R3-MYB and 1 4R-MYB genes.

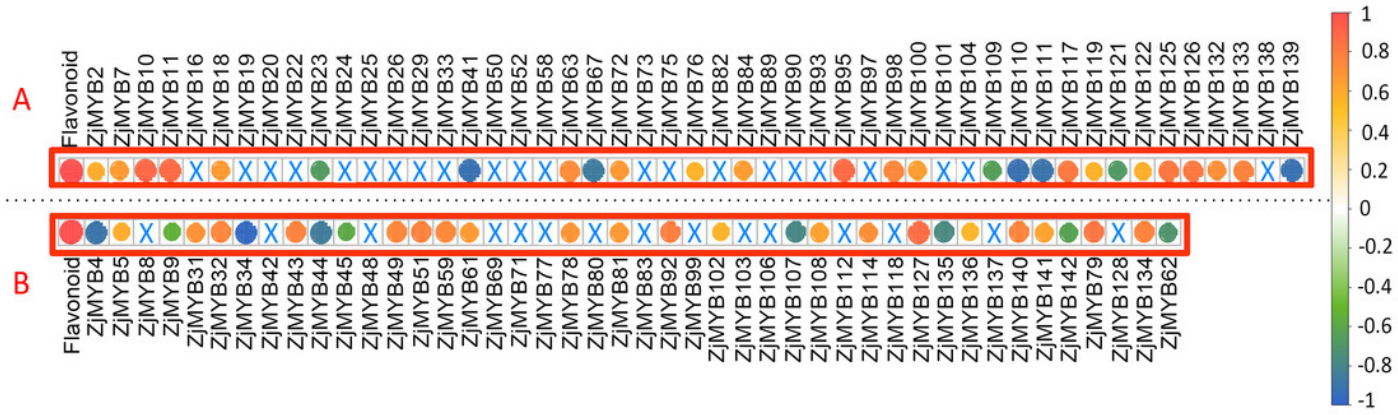
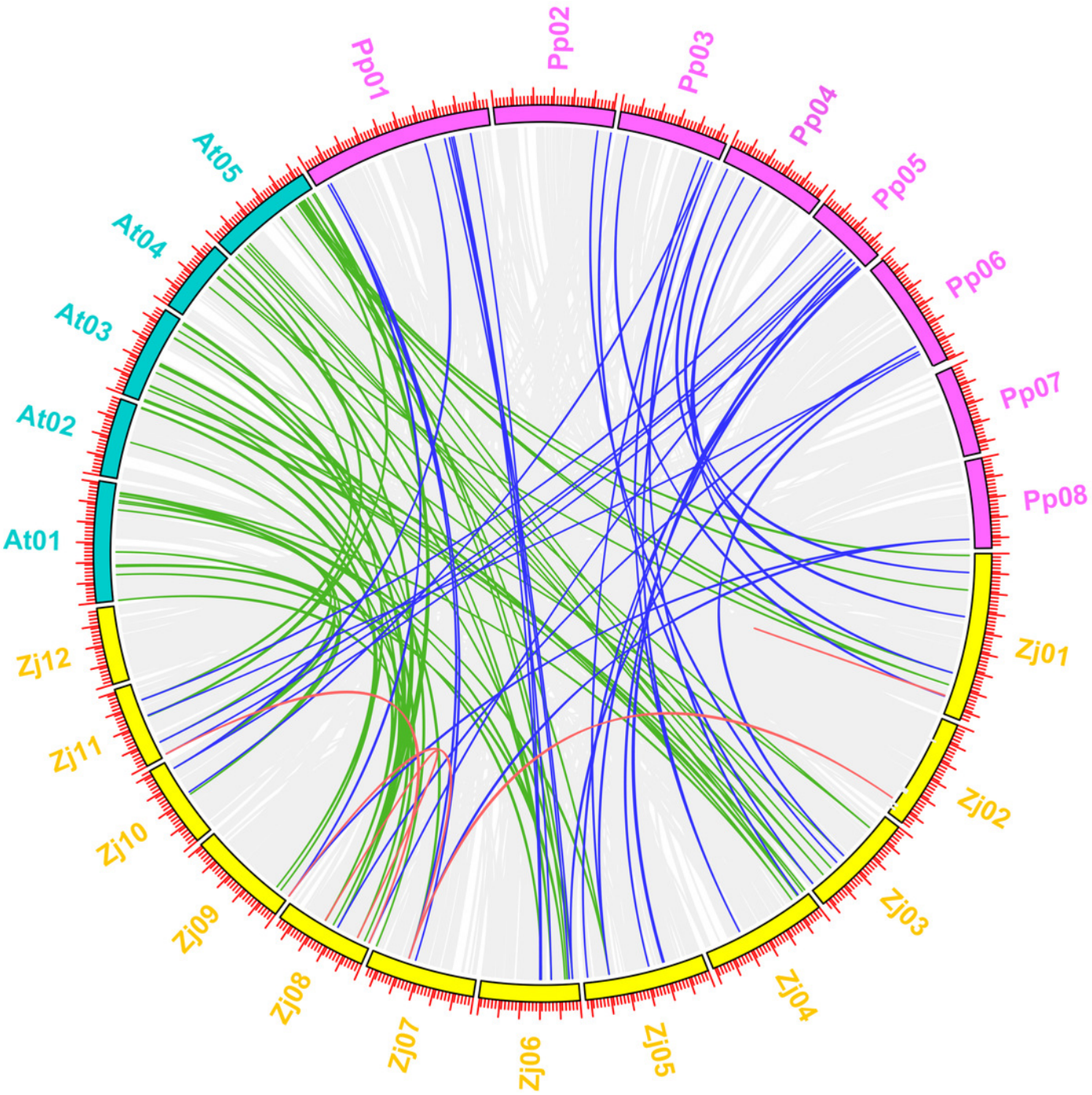


Figure 11

Synteny analysis of MYB genes between jujube and two other plant species (peach and *Arabidopsis*).

Twelve chromosomes of jujube (Zj01-Zj12) and five chromosomes of *Arabidopsis* (A01-A05), as well as eight chromosomes of peach (Pp01-Pp08), were mapped in different colors. The gray lines in the background indicate the collinear blocks within jujube and other plant genomes. The red lines connecting the MYB genes in jujube indicate paralogous MYB gene pairs. The green and blue lines connect the MYB genes in jujube and *Arabidopsis* and in jujube and peach that were orthologous, respectively.


Pp01
Pp02
Pp03
Pp04
Pp05
Pp06
Pp07
Pp08
Zj01
Zj02
Zj03
Zj04
Zj05
Zj06
Zj07
Zj08
Zj09
Zj10
Zj11
Zj12
At01
At02
At03
At04
At05