Genes Genom (2017) 39:831–842
DOI 10.1007/s13258-017-0544-5

Online ISSN 2092-9293
Print ISSN 1976-9571

RESEARCH ARTICLE

Genome-wide characterization and expression analysis of MYB transcription factors in *Lotus japonicas* and *Medicago truncatula*

Fang Wang[1] · **Xin Li**[2]

Received: 16 February 2017 / Accepted: 2 April 2017 / Published online: 9 April 2017

Abstract MYB is one of the most abundant families of plant transcription factors (TFs), and is involved in plant growth and development regulation, abiotic stress tolerance, hormone signal transduction, and disease resistance. Previously, few genome-wide assays have been conducted in *Lotus japonicas* and *Medicago truncatula*. Here, with the recent release of the genome sequences, we performed genome-wide study of MYB TFs in both, including family characterization, evolution and expression analysis. In this study, we totally identified 104 and 166 MYB genes in *L. japonicas* and *M. truncatula*, respectively. All these MYB genes were categorized into 261 R2R3-MYBs, 7 R1R2R3-MYBs and 2 atypical MYB genes. This result showed that the majority (96.7%) of them belong to the R2R3-MYB subfamily in both species. Based on phylogenetic analysis, MYB transcription factors were divided into 14 subgroups. MYB genes in four typical branches indicated species-specific expansions in *M. truncatula*. The exon numbers and conserved motifs associated with MYB domains were also characterized. Furthermore, gene expression analysis performed in various organs showed diverged expression profiles of most MYB genes, suggesting their potential roles in regulating organ growth and development in *L. japonicas* and *M. truncatula*.

Keywords Genome-wide study · MYB · *Lotus japonicas* · *Medicago truncatula* · Expression

Introduction

As the second most important plant families for agriculture, Legumes (Fabaceae) provide excellent materials for human food, animal feed, and industry use (Graham and Vance 2003). They account for one-third of crop production in the whole world, and play essential roles in symbiotic nitrogen fixation every year (Benedito et al. 2008). So far, many researchers have devoted to the study of breeding and genetic improvements of Legumes, such as *Glycine max, Lotus japonicas, Medicago truncatula*. In the past decades, with rapid development of microarray and genome sequencing technologies, the genome-wide research has become feasible. Currently, the genome sequencing of *L. japonicas* and *M. truncatula* were finished (Sato et al. 2008; Young et al. 2011). The gene expression atlases of both species were also established, respectively (Benedito et al. 2008; Verdier et al. 2013).

To better understand the regulation on growth and development of Legumes, we mainly aimed at the study on transcription factors (TFs), which temporarily and spatially control the expression of their target genes through binding upstream *cis*-elements (Jin et al. 2017). MYB is one of most abundant plant transcription factor (TF) families, and has been implicated in diverse plant-specific processes (Cedroni et al. 2003). The first plant MYB gene, isolated in maize, encoded a c-MYB-like TF, which was involved in the biosynthesis of anthocyanin

The original version of this article was revised.

Electronic supplementary material The online version of this article (doi:10.1007/s13258-017-0544-5) contains supplementary material, which is available to authorized users.

✉ Xin Li
lixin@njau.edu.cn

[1] Guangdong Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-Sen University, Guangzhou, China

[2] Laboratory Center of Life Sciences, College of Life Sciences, Nanjing Agricultural University, Nanjing, China

(Du et al. 2012). The most common type of plant MYB TFs is R2R3-MYB with two repeats (Jin et al. 2017). To date, substantial data about MYB transcription factors have been shown in both monocotyledonous and dicotyledonous plants (Feller et al. 2011).

MYB transcription factors were defined by the typical helix-turn-helix motifs (HTH) of their DNA-binding domains. Most MYB proteins in animals and plants have three (R1, R2, and R3) and two (R2 and R3) imperfect repeats, despite the fact that MYB genes containing both one and three repeats have been also found in plants (Braun and Grotewold 1999; Kranz et al. 2000). Thus, the MYB genes in *Arabidopsis thaliana* were classified into three types: (R1)R2R3_Myb, Myb_related and atypical_MYB (Yanhui et al. 2006). Notably, some groups also classified into three or four subfamilies according to the repeat numbers in the MYB domain (Stracke et al. 2001; Du et al. 2013; Li et al. 2016; Salih et al. 2016). Currently, genome-wide studies on R2R3-type MYB have been conducted in various plant species, such as *A. thaliana* (Stracke et al. 2001), diploid and polyploid cotton (Cedroni et al. 2003), *Zea mays* (Du et al. 2012), *Beta vulgaris* (Stracke et al. 2014), *Pyrus bretschneideri* (Li et al. 2016), *Jatropha curcas* (Peng et al. 2016), the tomato family Solanaceae (Gates et al. 2016). In spite of the extensive studies on R2R3-type MYB, the evolutionary history of MYB-related proteins in plants remains largely unknown (Du et al. 2013).

MYB TFs play important roles in plant growth and development, abiotic stress tolerance, hormone signal transduction and disease resistance (Jin and Martin 1999; Roy 2016). For example, expression profiles analysis in peanut (*Arachis hypogaea* L.) identified 30 MYB genes responsive to abiotic stress treatment (Chen et al. 2014). Overexpression of some MYB genes could lead to alternation of abiotic and biotic stress in tobacco (Li et al. 2014). R2R3-type MYB TFs were involved in secondary metabolism, such as phenylpropanoid metabolism (Jin and Martin 1999). MYB transcription factors, particularly R2R3-type MYB, play important roles in regulation of plant developmental processes, such as defense, cell shape, pigmentation, and root formation (Gates et al. 2016). Thus, MYB transcription factors may be related to organ development processes.

Here, we identified and characterized 104 and 166 *MYB* genes in *L. japonicas* and *M. truncatula*, respectively, most of which were R2R3-type MYB. Phylogenetic analysis indicated that MYB genes in *M. truncatula* underwent species-specific expansion. The expression analysis showed diverged expression profiles of most MYB genes in various organs, suggesting that they might be involved in plant organ growth and development.

Materials and methods

Identification of MYB genes

An extensive search was performed to identify MYB genes based on all protein sequences in the genomes of both *L. japonicas* (version: Lotus_r3.0) and *M. truncatula* (version: Mt4.0v1). All the coding sequences and protein sequences of both species were downloaded from the websites: ftp://ftp.kazusa.or.jp/pub/lotus/lotus_r3.0/ and http://www.jcvi.org/medicago/index.php. Based on the classification of MYB genes in *A. thaliana* (Yanhui et al. 2006), we collected 124 transcripts of R2R3-MYB genes. All the protein sequences were used for HMMER prediction of transcriptional factors under an *E* value with 1e-30, among which the classification of MYB genes were conducted according to that in *A. thaliana*. All the MYB genes were classified into three types: (R1)R2R3_Myb, Myb_related and atypical_MYB (Yanhui et al. 2006). Subsequently, all the coding sequences and protein sequences of the predicted MYB genes were extracted based on their gene/protein ID numbers.

The genomic locations of these MYB genes in *L. japonicas* and *M. truncatula* were extracted in the GFF files. Then these MYB genes were mapped to each chromosome of both species. We further predicted the MYB genes in both *L. japonicas* and *M. truncatula* subject to tandem duplication by MCScanX (Wang et al. 2012a).

Phylogenetic analysis

To characterize the evolutionary features of MYB proteins, we extracted all the predicted MYB protein sequences in both *L. japonicas* and *M. truncatula*. The protein sequences of *A. thaliana* MYB genes were obtained from the website TAIR (http://www.arabidopsis.org/). Using MEGA7.0 (Kumar et al. 2016), we constructed the phylogenetic trees of all MYB protein sequences in these three species. Both Neighbor-Joining (NJ) method and Maximum Likelihood (ML) method were employed for tree generation. In each method, bootstrap analysis with 1000 replicates was used.

Analysis of gene structures and conserved motifs

The GFF files of *L. japonicas* and *M. truncatula* were downloaded from the websites: ftp://ftp.kazusa.or.jp/pub/lotus/lotus_r3.0/ and http://www.jcvi.org/medicago/index.php. We then extracted the information of all the predicted MYB genes in both species, and used the Online Gene Structure Display Server (GSDS 2.0: http://gsds.cbi.pku.

edu.cn/) (Hu et al. 2015) to obtain the gene structures of four expanded gene clusters.

To analyze the sequence features of MYB repeats in R2R3-MYB proteins, we extracted the amino acid sequences of R2 and R3 repeats of all R2R3-MYB proteins in both *L. japonicas* and *M. truncatula*, and aligned them in each species by ClustalOmega (http://www.ebi.ac.uk/Tools/msa/clustalo/) using the default parameters. Using default settings, we employed the WebLogo (http://weblogo.berkeley.edu/logo.cgi) (Crooks et al. 2004) to create the sequence logos for R2 and R3 repeats from the multiple alignment files generated by ClustalOmega.

In order to identify conserved protein motifs in MYB TFs, we used the MEME software (version 4.11.2) (Bailey and Elkan 1994), with the parameter default settings.

Gene expression analysis

The microarray data of *L. japonicas* and *M. truncatula* were downloaded from the gene expression atlas website: http://ljgea.noble.org/v2/index.php and http://mtgea.noble.org/v3/index.php. The organ samples were collected in previous studies (Benedito et al. 2008; Verdier et al. 2013). In *L. japonicas*, the examined organs included Leaf, Nod21, Pod10d, Pod14d, Pod20, Pt (petiole), Root, Root0h, Seed10d, Seed12d, Seed14d, Seed16d, Seed20d, and Stem, while the *M. truncatula* organs included Flower, Leaf, Petiole, Pod, Root, Seed_10dap, Seed_12dap, Seed_16dap, Seed_20dap, Seed_24dap, Seed_36dap, Stem, and VegBud (vegetative bud). All these organs used for microarray experiments were harvested from plants under standard growth conditions. The microarray data were normalized as described by Benedito et al. (2008). The Z score was calculated according to previous studies (Benedito et al. 2008; Verdier et al. 2013). We determined the mean expression levels from three biological replicates of each organ. The expression data of MYB genes were then extracted in each species. The clustering analysis and the heat map generated were performed by R package heatmaps (https://cran.r-project.org/web/packages/pheatmap/index.html).

Results and discussion

Identification and classification of MYB transcription factors

To identify MYB genes in both the genomes of *L. japonicas* and *M. truncatula*, the HMMER prediction was performed using 198 MYB proteins in *A. thaliana* as queries, including 126 R2R3-MYB, 5 R1R2R3-MYB, 64 MYB-related, and 3 atypical MYB genes (Yanhui et al. 2006). As a result, 104 and 166 MYB genes were predicted in *L. japonicas* and *M. truncatula*, respectively. According to previous classification (Yanhui et al. 2006), there are 2 R1R2R3-MYB genes, 101 R2R3-MYB genes and 1 atypical MYB gene in *L. japonicas*. 5 R1R2R3-MYB genes, 160 R2R3-MYB genes and 1 atypical MYB gene were found in *M. truncatula*. In this study, the 104 and 166 MYB genes represented the MYB classification in *L. japonicas* and *M. truncatula*. Consistent with previous studies in plants (Li et al. 2016), the R2R3-MYB family is the most abundant MYB TFs in *L. japonicas* and *M. truncatula*, thus we further analyzed the R2R3-MYB family.

To map the MYB genes to chromosomes, the genomic locations of them were extracted from GFF files of all genes in *L. japonicas* and *M. truncatula*. All 104 MYB genes were mapped to 7 chromosomes (chromosome 0–7) (Fig. 1). 155 of 166 MYB genes (93.4%) were mapped to 8 chromosomes (chromosome 1–8) (Fig. 2), while 11 MYB genes were mapped to 9 unanchored scaffolds, including each MYB gene in 8 unanchored scaffolds and 3 MYB genes in one unanchored scaffold (Table 1). Each of 7 and 8 chromosomes contained R2R3-MYB genes, while the distributions of them were not even. Both the chromosome 0 in *L. japonicas* and chromosome 5 in *M. truncatula* contained 29 R2R3-MYB genes (Table 1), accounting for the most R2R3-MYB genes in one chromosome. Each of the 5th and 6th chromosomes contained 6 R2R3-MYB genes, and the 6th chromosome contained 11 R2R3-MYB genes (Table 1), representing the fewest numbers of R2R3-MYB genes in one chromosome. The uneven chromosome distribution of MYB genes may be due to uneven rates of gene duplication events.

Segmental duplication and tandem duplication were two of major mechanisms in origin and evolution of large gene families. From the genomic locations of MYB genes in *L. japonicas* and *M. truncatula*, we could conclude that both segmental and tandem duplication played important roles in shaping the evolution of MYB genes. Using MCScanX (Wang et al. 2012a), we obtained 9 and 13 MYB genes subject to tandem duplication (Supplementary Table 1). Interestingly, several gene clusters were subject to species-specific tandem duplication, because these tandem duplicated gene clusters were restricted to *L. japonicas* or *M. truncatula*.

MYB TFs have been studied in various plants (Jin et al. 2017), showing that the MYB gene numbers varied from 3 in *Helicosporidium* to 489 in *Brassica napus*. Even in Legumes, the MYB TFs vary greatly, from 117 in *Vigna radiate* to 430 in *Glycine max* (Jin et al. 2017). This indicated that expansion of MYB TFs occurred in the evolution of Legumes, particularly in *G. max*. In this study, we identified 104 and 166 MYB TFs in *L. japonicas* and *M. truncatula*, suggesting MYB expansion in *M. truncatula*.

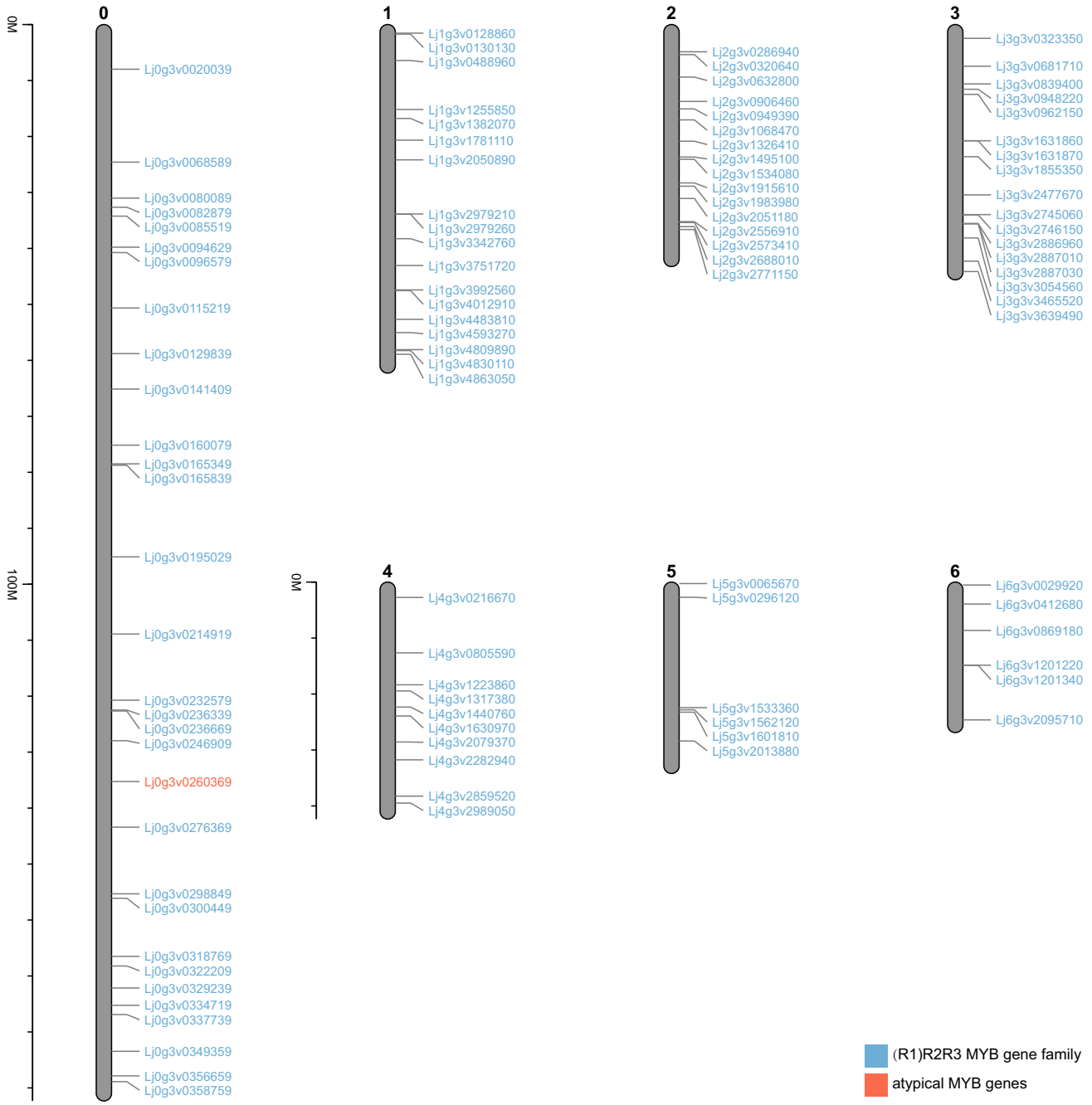

Fig. 1 The chromosome location of MYB genes in *L. japonicas*

Expansion of MYB transcription factors in *M. truncatula*

To evaluate the evolutionary significance of the MYB genes, we performed phylogenetic analysis of *L. japonicas, M. truncatula* and *A. thaliana* MYB proteins using Neighbor-Joining (NJ) method and Maximum Likelihood (ML) method (Fig. 3 and Supplementary Fig. 1). Generally, the topology of them was similar. All the MYB genes of three species in the NJ tree could be divided to 14 subgroups (Fig. 3). Consistent with previous studies (Wang et al. 2015; Salih et al. 2016), the bootstrap values for some subgroups of the NJ tree were low due to relatively large number of gene sequences.

In order to validate the NJ tree, the phylogenetic tree of MYB genes was also reconstructed with ML method. The results here showed that the trees constructed by both methods mentioned above were mainly consistent with

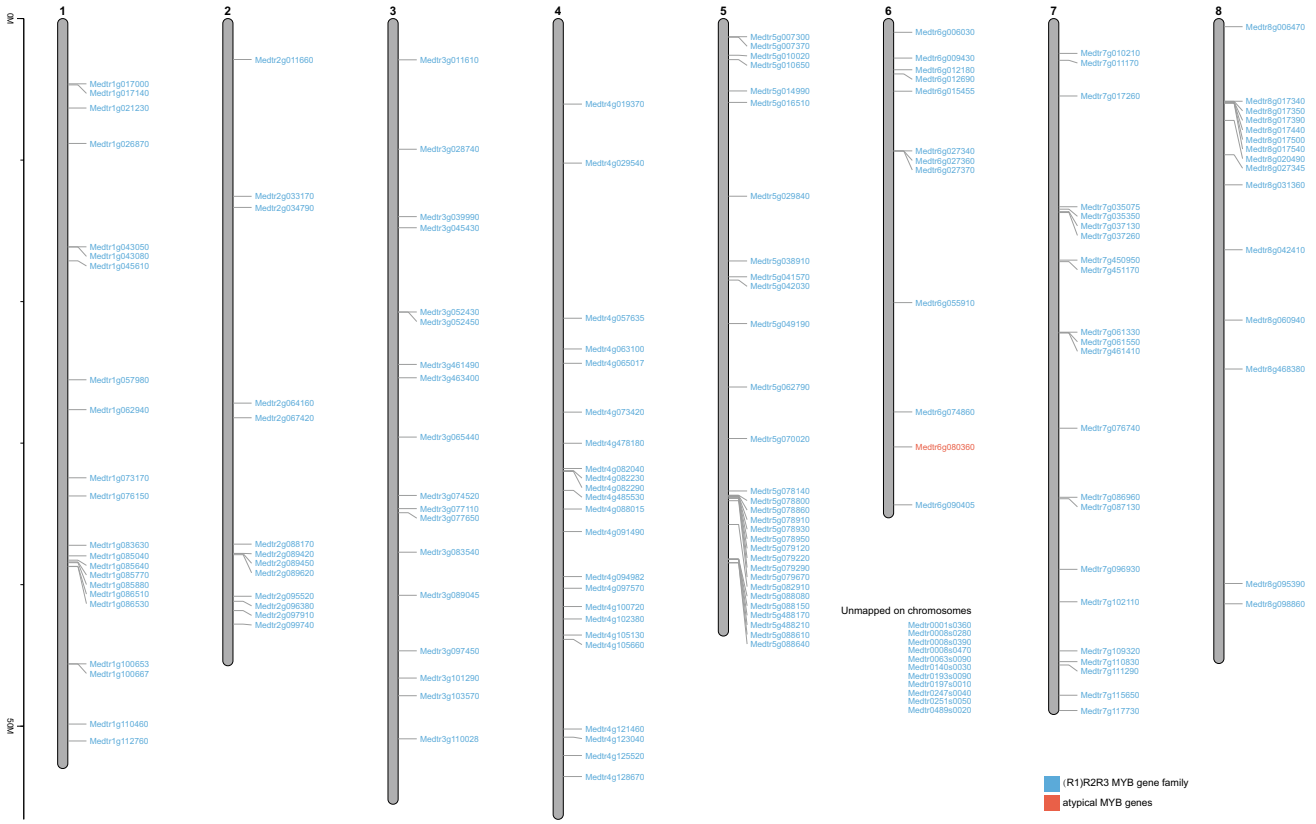

Fig. 2 The chromosome location of MYB genes in *M. truncatula*

each other (Fig. 3 and Supplementary Fig. 1). Thus, the phylogenetic trees of MYB genes were reliable.

Subgroup 14 was the oldest MYB subgroup, while the subgroup 1 was the youngest one. Particularly, the subgroup 1, 5, 6, 8, 9 and 13 contained more MYB genes in *M. truncatula* than that in *L. japonicas* (Fig. 3), supporting the species-specific expansion of R2R3-MYB genes in *M. truncatula*. As shown in Supplementary Table 2, four gene clusters with 36 MYB genes were expanded in both NJ and MP trees, supporting the species-specific expansion of R2R3-MYB genes in *M. truncatula*. No similar expansion events with at least seven genes were observed in *L. japonicas* (Fig. 3 and Supplementary Fig. 1).

Additionally, the MYB genes in subgroup 7 were specific to *L. japonicas* and *M. truncatula*, but absent in *A. thaliana* (Fig. 3 and Supplementary Fig. 1). Further sequence analysis showed that the part of MYB proteins in the subgroup 7 were homologous to other proteins in *G. max, A. thaliana, Z. mays* and *O. sativa*. Therefore, these MYB TFs may not be specific to Legumes. However, other parts of them have no homology in *G. max, A. thaliana, Z. mays* or *O. sativa*, suggesting that they might only emerge in these two species through partial gene duplication. Therefore, *M. truncatula* may evolve novel R2R3-MYB genes to regulate gene expression.

Features of gene structures and conserved domains in MYB genes

To investigate the gene structure features, we collected the exon numbers of all 270 MYB genes in both *L. japonicas* and *M. truncatula*. Most of the R2R3-MYB genes in *L. japonicas* (95.0%) and *M. truncatula* (98.1%) contained between 1 and 12 introns (Fig. 4). Most R2R3-MYB genes in both *L. japonicas* (67.3%) and *M. truncatula* (75.6%) contained two introns, followed by that containing one intron (18.8% in *L. japonicas* and 13.8% in *M. truncatula*) (Fig. 4 and Supplementary Fig. 2). Only a small portion of R2R3-MYB genes were intronless. This result was similar to that described in *Arabidopsis, Vitis vinifera, Eucalyptus grandis* and *Gossypium hirsutum* (Matus et al. 2008; Soler et al. 2015; Salih et al. 2016). Notably, all R1R2R3-MYB genes and atypical MYB genes in both species contained no less than seven exons (Supplementary Table 3).

We further studied the variation within the conserved motifs of R2R3-MYB genes in both *L. japonicas* and *M. truncatula*, using the WebLogo program (Crooks et al. 2004). The R2 motifs showed similar amino acid compositions between *L. japonicas* and *M. truncatula* (Fig. 4), suggesting similar structures and functions between both species. A similar result of R3 motifs was observed

Table 1 Chromosome distributions of three MYB subfamilies in *L. japonicas* and *M. truncatula*

Species	Chromosome	Transcription factor types		
		R1R2R3 MYB	R2R3 MYB	Atypical MYB
L. japonicas	chr0	1	29	1
	chr1	1	17	
	chr2		16	
	chr3		17	
	chr4		10	
	chr5		6	
	chr6		6	
M. truncatula	chr1	1	21	
	chr2		13	
	chr3	1	17	
	chr4		23	
	chr5	1	29	
	chr6		11	1
	chr7	2	20	
	chr8		15	
	scaffold0001		1	
	scaffold0008		3	
	scaffold0063		1	
	scaffold0140		1	
	scaffold0193		1	
	scaffold0197		1	
	scaffold0247		1	
	scaffold0251		1	
	scaffold0489		1	

(Fig. 4). However, the amino acid compositions between R2 and R3 in the same species were different (Fig. 4), suggesting functional divergence between them. Although the functional divergence of these R2R3-MYB genes remained to be identified, they were thought to play important roles in regulation of gene expression and functional diversification in plants.

Additionally, MEME results showed that most of MYB TFs contained at least three motifs, but the motif sequences were not the same between subgroups, also supporting sequence and even functional divergence of these MYB TFs (Supplementary Fig. 2 and Supplementary Table 4). Several MYB genes lost some motifs in both *L. japonicas* and *M. truncatula*. *Lj2g3v1534080* (subgroup 2) and *Lj4g3v1630970* (subgroup 4) lost the 3rd motif, *Lj6g3v2095710* (subgroup 5), *Medtr5g007300* (subgroup 5), *Lj0g3v0195029* (subgroup 6), and *Lj2g3v0320640* (subgroup 6) lost the 2nd motif. In subgroup 8, Lj0g3v0334719 and Lj3g3v3054560 lack the 1st motif. These observations suggested structural diversifications of these MYB TFs in *L. japonicas* and *M. truncatula*.

Divergent expression of MYB genes

In *L. japonicas*, 84 of 104 MYB genes (~80.8%), represented by 120 probe sets, were expressed in at least one of all the investigated organs in this study, while 88 of 166 MYB genes (~53.0%), represented by 99 probe sets, were could be detected in the microarrays of *M. truncatula* organs, suggesting more than 50% of the MYB genes could be detected in the microarray system of previous studies (Benedito et al. 2008; Verdier et al. 2013). 80.3 and 86.1% of all genes in *L. japonicas* and *M. truncatula* were expressed in one or more organs (Benedito et al. 2008; Verdier et al. 2013). Then the average expression values of MYB genes and all genes in both species were calculated. As shown in Table 2, the expressed MYB genes in both species exhibited lower average expression levels than all genes ($P < 0.05$, *t* test).

Analysis of organ-specific genes may provide insight into specialized organ processes, including biochemical, physiological, developmental and other processes (Verdier et al. 2013). According to this previous study, we calculated Z scores for each probe set to identify organ-specific genes.

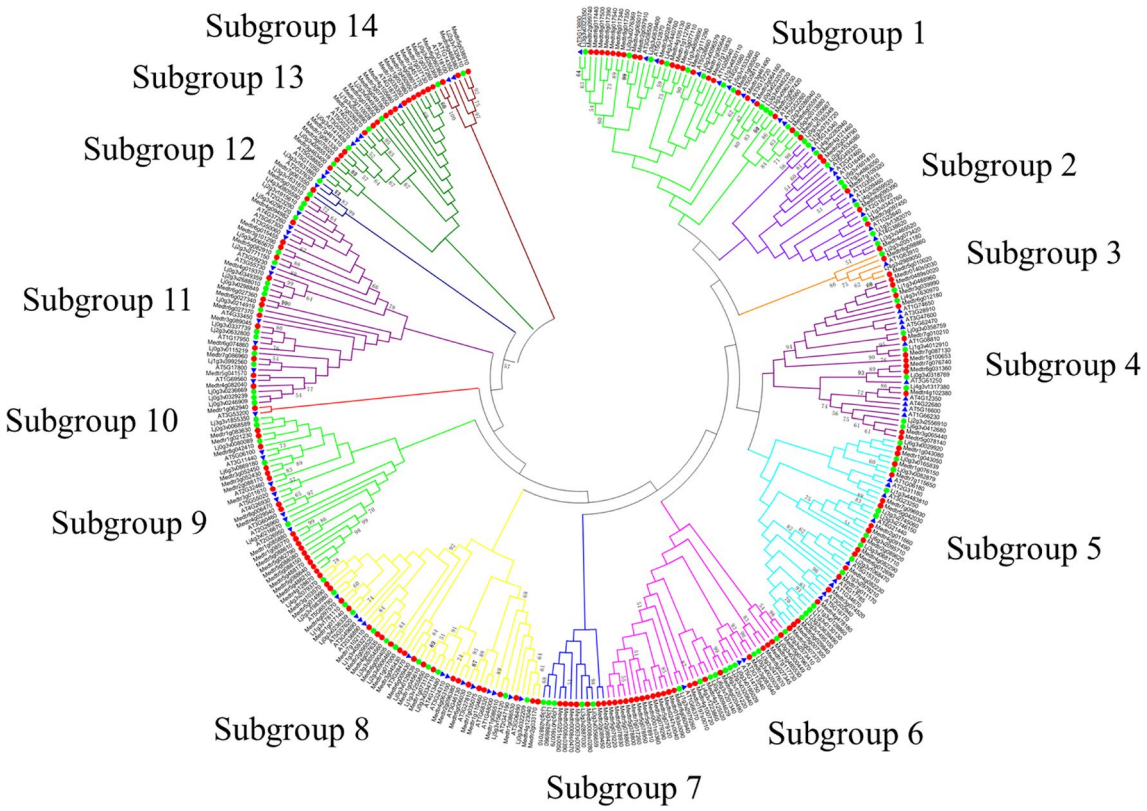

Fig. 3 The Neighbor-Joining (NJ) phylogenetic tress of all MYB genes in *A. thaliana, L. japonicas* and *M. truncatula. filled circle* denotes MYB genes in *L. japonicas* and *M. truncatula*, in which *red* denote MYB genes in *M. truncatula*, while *green* denote MYB genes in *L. japonicas; filled triangle* denotes MYB genes in *A. thaliana*

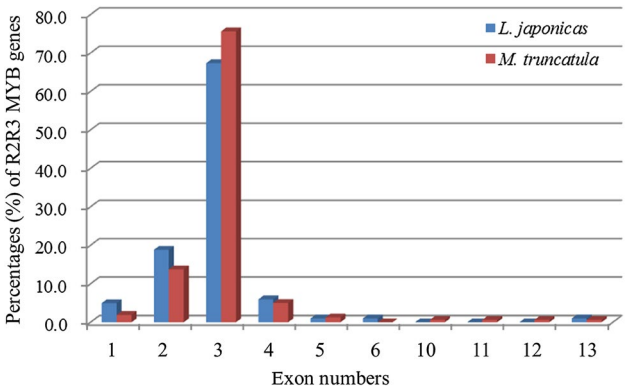

Fig. 4 The exon numbers of all MYB genes in *L. japonicas* and *M. truncatula*

Similarly, a minimum Z-score of 2.85 and a minimum normalized expression value >100 were used as threshold values. Interestingly, all of MYB genes in *L. japonicas* and *M. truncatula* have Z-scores less than 2.85 (Fig. 5); therefore, they were expressed in at least two organs.

We further analyzed the expression pattern in these organs. Some MYB genes within the same subgroup exhibit similar expression patterns in these organs of *L. japonicas* and *M. truncatula*. As shown in Supplementary Tables 5, 9 and 12 gene clusters within the same subgroups were observed to exhibit similar expression profiles in *L. japonicas* and *M. truncatula*. For example, *Lj2g3v0320640* and *Lj6g3v1201340* in subgroup 6 were higher expressed in pod than in other organs (Fig. 6). *Lj0g3v0214919* and *Lj0g3v0115219* in subgroup 11 exhibited similar expression profile (Fig. 6). The expression levels of *Medtr3g083540* and *Medtr8g020490* in subgroup 6 were higher in seeds than in other organs (Fig. 7). *Medtr0140s0030* and *Medtr0489s0020* in subgroup 4 were highly expressed in leaf, Vegetative-Bud and pod (Fig. 7). In spite of these observations, most of MYB genes within the same subgroups showed different expression patterns (Figs. 6, 7), indicating expression divergence and even functional divergence. As mentioned above, these MYB genes might arise through segmental and tandem duplication. Previously, many studies have showed expression divergence after gene duplication was a general pattern in the course of plant evolution (Casneuf et al. 2006; Li et al. 2009; Wang et al. 2012b). Our data here also support this expression pattern.

Table 2 The expression of all MYB genes and all genes expressed in *L. japonicas* and *M. truncatula*

Species	All genes			MYB genes			
	Gene number	Probe number	Average expression value	Gene number	Probe number	Average expression value	*t*-test *p*-value
L. japonicas	24,184 (39,734)	61,459	221.561	84 (104)	120	69.483	2.2e-16
M. truncatula	23,302 (50,894)	61,278	246.017	88 (166)	99	114.377	2.2e-16

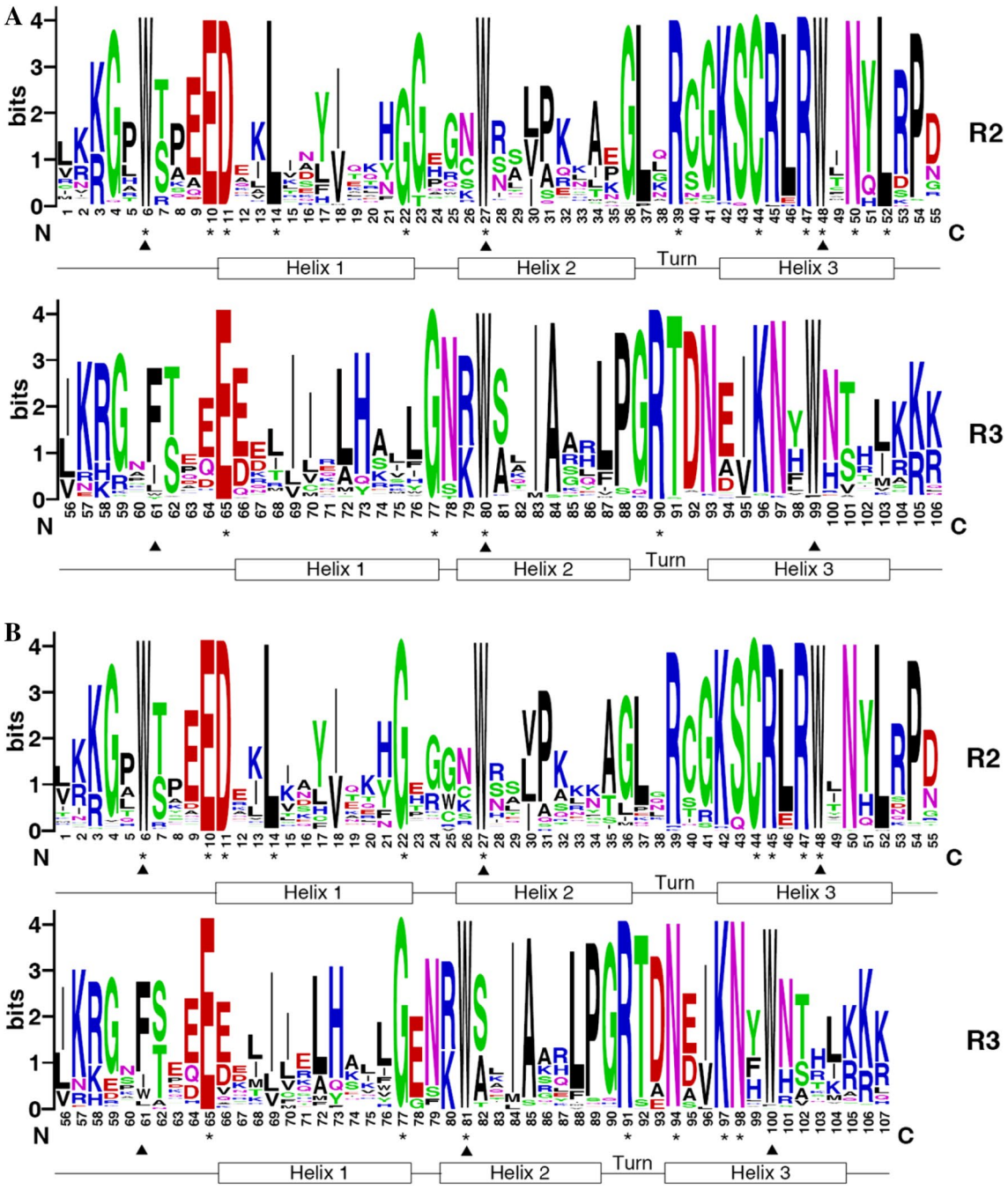

Fig. 5 The conserved motifs of the R2R3 MYB genes in *L. japonicas* and *M. truncatula*

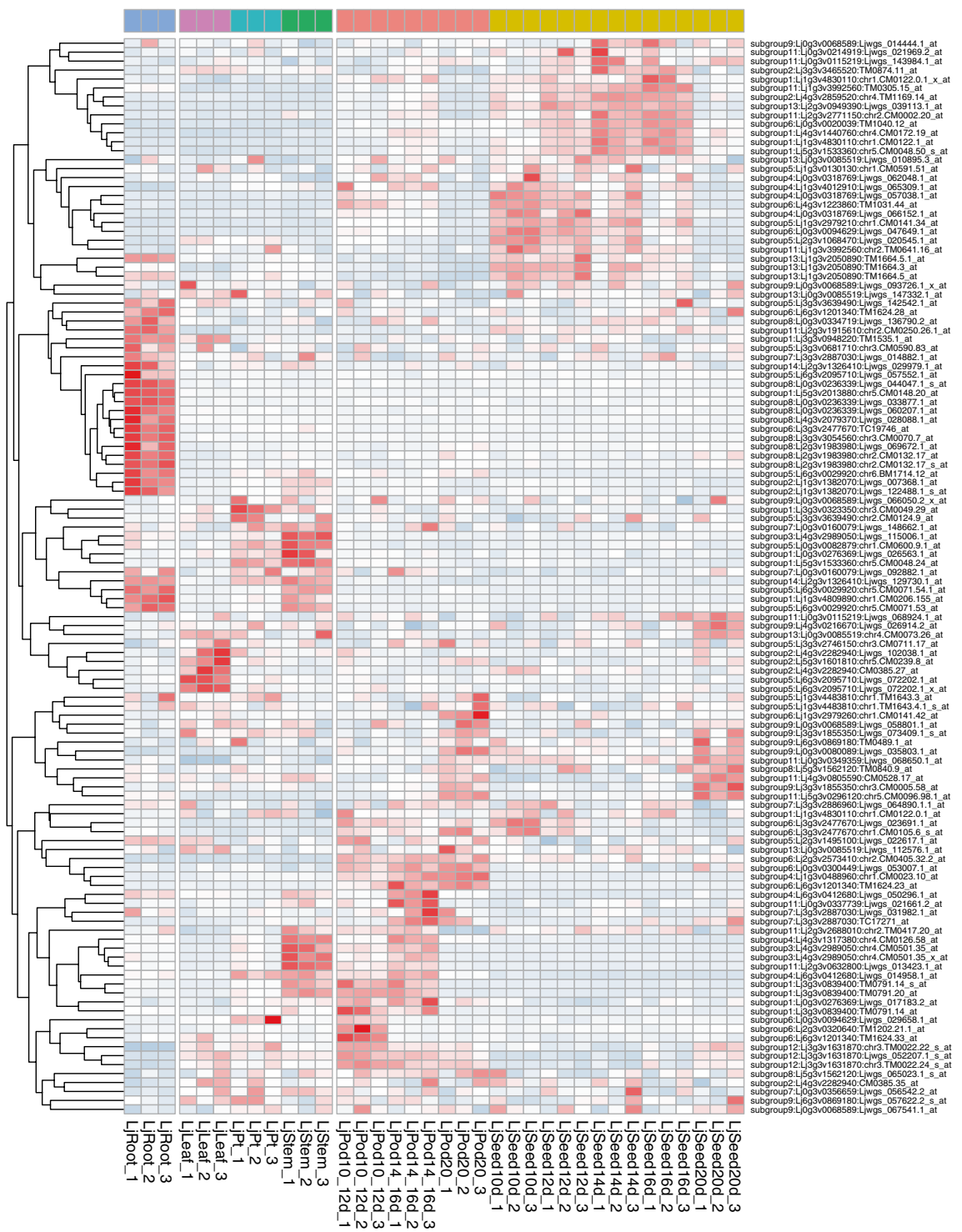

Fig. 6 The heatmap of all MYB genes in *L. japonicas*

Particularly, only a small portion (25.0%) of the 4 expanded 36 MYB genes in 4 subgroups (Supplementary Table 2) has expression data in *M. truncatula*, suggesting that most of them may be lowly expressed. These species-specific expanded MYB genes might arise through recent duplication events. Newly duplication genes were usually functional redundant, thus the MYB genes here may reduce this redundancy through down-regulating expression levels.

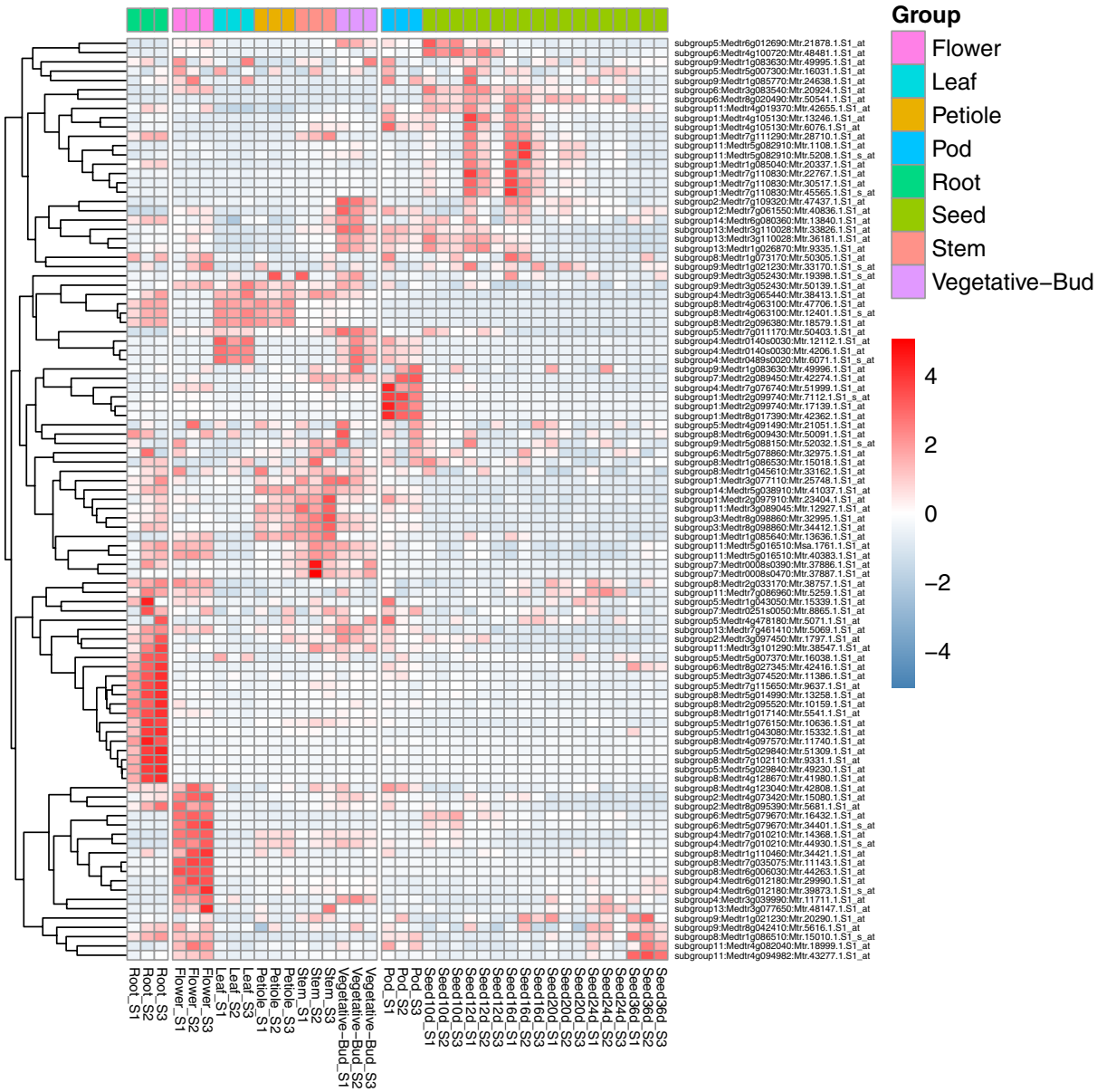

Fig. 7 The heatmap of all MYB genes in *M. truncatula*

Roles of MYB TFs in plant development

To characterize the functions of MYB TFs in *L. japonicas* and *M. truncatula*, we collected functional assignments of 20 MYB TFs in *A. thaliana* within 14 subgroups, according to previous study (Stracke et al. 2014). As shown in Supplementary Table 6, MYB TFs within these 11 subgroups were involved in cell cycle, defense, development, differentiation and metabolism. We also collected putative functions of 22 MYB TFs in *A. thaliana* within 14 subgroups as summarized in a previous report about Chinese White Pear (*P. bretschneideri*) (Supplementary Table 7) (Li et al. 2016). The MYB TFs within the same subgroups may share similar functions, thus we could obtain functional clues of MYB TFs in *L. japonicas* and *M. truncatula*. For example, the MYB TFs within Subgroup 1 was related to development and metabolism, such as shoot morphogenesis and leaf patterning, root development, and lignin biosynthesis (Supplementary Tables 6 and 7). Taken the summary together, MYB TFs within Subgroup 1, 3, 4, 5, 6, 8, 9, 11, 12 and 13 were related to development (Supplementary Tables 6 and 7), covering most of the 14 subgroups. Thus, most of the MYB TFs were involved in plant development.

Conclusions

In summary, the present study identified 270 MYB genes in *L. japonicas* and *M. truncatula* according to the *A. thaliana* genome. The phylogenetic relationships between subfamilies, conserved motifs, expression patterns in different organs, were surveyed in detail. Our results provide better understanding about the molecular basis of MYB genes on the organ growth and development of *L. japonicas* and *M. truncatula*, which allow us to obtain a better platform for interesting MYB gene research in the future.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 31200179) and the Ministry of Agriculture of China for Transgenic Research (Grant Nos. 2014ZX0800943B).

Compliance with ethical standards

Conflict of interest Fang Wang and Xin Li declare that they have no conflict of interest.

Research involving human and animal participants This article does not contain any studies with human subjects or animals performed by any of the authors.

References

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28–36

Benedito VA, Torres-Jerez I, Murray JD, Andriankaja A, Allen S, Kakar K, Wandrey M, Verdier J, Zuber H, Ott T et al (2008) A gene expression atlas of the model legume *Medicago truncatula*. Plant J 55(3):504–513

Braun EL, Grotewold E (1999) Newly discovered plant c-myb-like genes rewrite the evolution of the plant myb gene family. Plant Physiol 121(1):21–24

Casneuf T, De Bodt S, Raes J, Maere S, Van de Peer Y (2006) Nonrandom divergence of gene expression following gene and genome duplications in the flowering plant *Arabidopsis thaliana*. Genome Biol 7(2):R13

Cedroni ML, Cronn RC, Adams KL, Wilkins TA, Wendel JF (2003) Evolution and expression of MYB genes in diploid and polyploid cotton. Plant Mol Biol 51(3):313–325

Chen N, Yang Q, Pan L, Chi X, Chen M, Hu D, Yang Z, Wang T, Wang M, Yu S (2014) Identification of 30 MYB transcription factor genes and analysis of their expression during abiotic stress in peanut (*Arachis hypogaea* L.). Gene 533(1):332–345

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14(6):1188–1190

Du H, Feng BR, Yang SS, Huang YB, Tang YX (2012) The R2R3-MYB transcription factor gene family in maize. PLoS ONE 7(6):e37463

Du H, Wang YB, Xie Y, Liang Z, Jiang SJ, Zhang SS, Huang YB, Tang YX (2013) Genome-wide identification and evolutionary and expression analyses of MYB-related genes in land plants. DNA Res 20(5):437–448

Feller A, Machemer K, Braun EL, Grotewold E (2011) Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J 66(1):94–116

Gates DJ, Strickler SR, Mueller LA, Olson BJ, Smith SD (2016) Diversification of R2R3-MYB transcription factors in the tomato family Solanaceae. J Mol Evol 83(1–2):26–37

Graham PH, Vance CP (2003) Legumes: importance and constraints to greater use. Plant Physiol 131(3):872–877

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31(8):1296–1297

Jin H, Martin C (1999) Multifunctionality and diversity within the plant MYB-gene family. Plant Mol Biol 41(5):577–585

Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G (2017) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res 45(D1):D1040–D1045

Kranz H, Scholz K, Weisshaar B (2000) c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. Plant J 21(2):231–235

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874

Li Z, Zhang H, Ge S, Gu X, Gao G, Luo J (2009) Expression pattern divergence of duplicated genes in rice. BMC Bioinform 10(Suppl 6):S8

Li JB, Luan YS, Yin YL (2014) SpMYB overexpression in tobacco plants leads to altered abiotic and biotic stress responses. Gene 547(1):145–151

Li X, Xue C, Li J, Qiao X, Li L, Yu L, Huang Y, Wu J (2016) Genome-wide identification, evolution and functional divergence of MYB transcription factors in Chinese white pear (*Pyrus bretschneideri*). Plant Cell Physiol 57(4):824–847

Matus JT, Aquea F, Arce-Johnson P (2008) Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across *Vitis* and *Arabidopsis* genomes. BMC Plant Biol 8:83

Peng X, Liu H, Wang D, Shen S (2016) Genome-wide identification of the *Jatropha curcas* MYB family and functional analysis of the abiotic stress responsive gene *JcMYB2*. BMC Genom 17:251

Roy S (2016) Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. Plant Signal Behav 11(1):e1117723

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(1):129

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K et al (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15(4):227–239

Soler M, Camargo EL, Carocha V, Cassan-Wang H, San Clemente H, Savelli B, Hefer CA, Paiva JA, Myburg AA, Grima-Pettenati J (2015) The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. New Phytol 206(4):1364–1377

Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. Curr Opin Plant Biol 4(5):447–456

Stracke R, Holtgrawe D, Schneider J, Pucker B, Sorensen TR, Weisshaar B (2014) Genome-wide identification and characterisation of R2R3-MYB genes in sugar beet (*Beta vulgaris*). BMC Plant Biol 14:249

Verdier J, Torres-Jerez I, Wang M, Andriankaja A, Allen SN, He J, Tang Y, Murray JD, Udvardi MK (2013) Establishment of the *Lotus japonicus* Gene Expression Atlas (LjGEA) and its use to explore legume seed maturation. Plant J 74(2):351–362

Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H et al (2012a) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40(7):e49

Wang Y, Wang X, Paterson AH (2012b) Genome and gene duplications and gene expression divergence: a view from plants. Ann N Y Acad Sci 1256:1–14

Wang Z, Tang J, Hu R, Wu P, Hou XL, Song XM, Xiong AS (2015) Genome-wide analysis of the R2R3-MYB transcription factor genes in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) reveals their stress and hormone responsive patterns. BMC Genom 16:17

Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q et al (2006) The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol Biol 60(1):107–124

Young ND, Debelle F, Oldroyd GE, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, Gouzy J, Schoof H et al (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature 480(7378):520–524