BMC Genomics

RESEARCH Open Access

Genome-wide identification, expression analysis of the *R2R3-MYB* gene family and their potential roles under cold stress in *Prunus sibirica*

Xin Zhao[1,2], Shipeng Wang[1,2], Hongrui Zhang[1,2], Shengjun Dong[1,2], Jianhua Chen[1,2], Yongqiang Sun[1,2], Yueyuan Zhang[1] and Quangang Liu[1,2*]

Abstract

Background The R2R3-MYB transcription factors in plants participate in various physiological and biochemical processes and responds to various external stimuli. *Prunus sibirica* (known as Siberian apricot) is a drupe tree species that produces extremely high nutritional value kernels. However, it is susceptiblility to frost damage during the flowering period, results in a marked reduction in kernel yield.

Results In this study, the *MYB* gene family of *P. sibirica* (*PsMYB*) was systematically analyzed, and 116 *R2R3-MYB* genes that were distributed unevenly over eight chromosomes were ultimately screened. Phylogenetic analysis divided these 116 genes into 30 subgroups. We discovered that 37 *PsMYBs* had cold stress-responsive promoters, and six *PsMYBs* were annotated to be associated with cold response. Intraspecific homology analysis identified segmental duplication as the primary gene amplification mechanism, and homology analysis of the *PsMYB* genes with those of five other species revealed phylogenetic relationships with *Rosaceae* species. Protein interaction studies revealed collaborative regulation of the PsMYB proteins with *Arabidopsis* protein, and transcriptome analysis identified *PsMYB* genes that were highly expressed at low temperatures. Additionally, the expression levels of 22 *PsMYBs* in different tissue parts of *P. sibirica* and under different low-temperature stress conditions were evaluated using quantitative real-time PCR, with the results verifying that *PsMYBs* are specifically expressed in different plant parts and may be involved in the growth and development of *P. sibirica* species. Genes upregulated after exposure to low-temperature stress and likely involved in cold response were identified.

Conclusion This study lays a foundation for understanding the molecular biology of *PsMYBs* in *P. sibirica* and provides a theoretical basis for the future study of transgenic lines with cold resistance during the flowering period of this tree.

Keywords MYB transcription factor, Low-temperature stress, Expression patterns, Siberian apricot

*Correspondence:
Quangang Liu
liuquangang007@126.com
[1]College of Forestry, Shenyang Agricultural University, Shenyang 110866, China
[2]Key Laboratory for Silviculture of Liaoning Province, Shenyang Agricultural University, Shenyang 110866, China

Background

The family of avian myeloblastosis proto-oncogene (MYB) transcription factors (TFs) is one of the largest and most abundant of its kind in plants [1]. Each N terminus of MYB family proteins has a highly conserved DNA-binding domain (DBD) consisting of 1–4 incomplete tandem repeats (R), each comprising 50–53 conserved amino acid (aa) residues [2]. MYB TFs can be divided into four main types according to the number and position of the R sequences: 2R (R2R3-MYB), 3R (R1R2R3-MYB), 4R (R1R2R2R1/2-MYB), and 1R-MYB (MYB-related proteins) [3, 4]. Each R forms a helix–turn–helix (HTH) structural motif with three conserved tryptophan residues (W) [5]. The first MYB gene ever identified was the v-Myb oncogene of the avian myeloblastosis virus [6]. The first MYB TFs discovered in plants was from the maize plant *Zea mays* (*ZmMYBC*) and was proven to be a product of anthocyanin biosynthesis [7]. The 2R and 3R sequences are also crucial for identifying the MYB family menbers [8]. *R2R3-MYB* is the most common subfamily of the plant *MYB* gene family and the most studied [9, 10]. Many *R2R3-MYB* genes have been identified across a wide range of plant species; for example, 80 in *Spinacia oleracea* [11], 126 in *Arabidopsis thaliana* [12], 196 in *Populus trichocarpa* [13], 99 in rice (*Oryza*) species [14], 176 in *Stevia rebaudiana* [15], and 216 in *Salix* [16]. However, the genome-wide identification of the *R2R3-MYB* gene family in the Siberian apricot (*Prunus sibirica*) has not yet been reported.

In plants, R2R3-MYB TFs participate in various developmental processes as well as biological and abiotic regulation. For example, 12 *R2R3-MYB* candidate genes related to color development were identified in *Cocos nucifera* [17]. *LoMYB33* was found to be involved in pollen development in lilies (*Lilium oriental* hybrids) [18]. The activity of MYB TFs under salt stress has been investigated using bioinformatics technology bioinformatics technology and *Petunia* as a model species [19]. The *HbR2R3-MYB* gene members of the rubber tree (*Hevea brasiliensis*) exhibited different levels of expression under powdery mildew infection and ethylene treatment, implying the diversity of their functions [20]. The pepper (*Capsicum annuum*) gene *CaMYB306* plays a negative role in cold response by regulating the reactive oxygen species (ROS) system and inhibiting the transcriptional activity of calcineurin B-like interacting protein kinase gene 13 (*CaCIPK13*) [21]. Transfer of the *FvMYB82* gene of *Fragaria vesca* to *A. thaliana* substantially increased the proline and chlorophyll levels and superoxide dismutase (SOD, EC 1.15.1.1), peroxidase (POD, EC 1.11.1.7), and catalase (CAT, EC 1.11.1.6) activities of the transgenic lines, which significantly improved the tolerance of *A. thaliana* to salt and cold stress [22]. The *LpMYB* gene members of *Lilium pumilum* are upregulated to varying degrees under low-temperature, drought, and high-salt stress [23]. The *PhMYB62* of *Pyrus hopeiensis* flowers can bind to POD gene promoter, enhance the detoxification of ROS in plant cells, and increase the cold resistance [24].

P. sibirica is a drupe tree species belonging to the *Rosaceae* family of trees and shrubs found in East Asia and Siberia has both ecological and economic value. The essential amino acid (EAA) content of its kernels is higher than that of other similar almonds [25]. However, *P. sibirica* is an early flowering tree species that is highly susceptible to damage from low temperatures and late frost, which results in a serious decrease in kernel yield and constrains the sustainable development of the almond industry [26]. To address the reduced yield and poor quality of *P. sibirica* kernels caused by frost damage during the flowering period, new economic varieties with high cold resistance during the flowering period need to be developed. In this regard, identifying the genes related to frost resistance in the flower organs is of great significance for attaining novel transgenic lines.

In this study, we used bioinformatics technology to characterize and analyze the *R2R3-MYB* gene family members in *P. sibirica* and identified 116 *R2R3-MYB* genes (*PsMYBs*). A comprehensive predictive analysis was performed on select *PsMYBs* on the basis of conserved domains, phylogenetic trees, transcriptomes, and other aspects. Quantitative Real-time PCR (qRT-PCR) was used to evaluate the expression patterns of specific *PsMYB* genes in different tissues under low-temperature stress. This study aimed to comprehensively understand the R2R3-MYB TF family and predict their functions, laying the theoretical foundation for functional studies of this family and investigating the molecular mechanism of flower organ resistance to freezing in *P. sibirica*. The findings of this study will contribute to improving the yield of *P. sibirica* kernels and provide support for enriching germplasm resources through genetic engineering.

Results

Identification of the *R2R3-MYB* gene family in *P. sibirica*

In this study, 126 *A. thaliana* R2R3-MYB proteins, consensus protein sequences of the MYB-binding domain, and a Hidden Markov Model (HMM) profile (PF00249) were used to retrieve the *P. sibirica R2R3-MYB* genome. In total, 116 *R2R3-MYB* genes were confirmed in the *P. sibirica* genome and designated as *PsMYB1-PsMYB116* on the basis of their chromosomal locations.

The physicochemical properties of the 116 PsMYB proteins were analyzed, and their detailed characteristics, including number of amino acids (aa), molecular weight (MW), isoelectric point (pI), instability index (II), aliphatic index (AI), and grand average hydropathicity (GRAVY), were recorded. Among the 116 TFs,

PsMYB85 was identified as the smallest protein with only 148 amino acids, whereas PsMYB20 was the largest with 1038 amino acids. The relative molecular weights of the proteins ranged from 17.1 kDa (PsMYB85) to 113.8 kDa (PsMYB20). The theoretical pI ranges values ranged from 4.84 (PsMYB64) to 9.98 (PsMYB85). Moreover, 58 of the 116 proteins were acidic (pI < 7) and 58were alkaline (pI > 7). All 116 TFs had negative (GRAVY < 0), indicating that all were hydrophilic proteins, with PsMYB116 exhibiting the strongest hydrophilicity (–1.117). According to the instability index values, only four TFs (PsMYB43, PsMYB78, PsMYB85, and PsMYB100) were stable (II < 40). The AI values of PsMYB5 and PsMYB116 were 80.21 and 50.82. respectively. Prediction of the subcellular localization of the TFs localized 112 of them to the nucleus. Furthermore, PsMYB5, PsMYB19, and PsMYB85 were located in the chloroplasts, whereas PsMYB87 was localized in the outer chloroplast membrane (Table 1).

Multiple sequence alignment, gene structure, and motif conservation

Multiple amino acid sequence alignment revealed that the proteins had significant R2R3 structures, and the highly conserved repetitive tryptophan "W" was labeled in the WebLogo diagram. The R2 repeat sequence contained three tryptophan residues (W). In the R3 repeat sequence, the first tryptophan residue was easily replaced by other hydrophobic amino acids, such as phenylalanine (F), isoleucine (I), and leucine (L), while the second and third tryptophan residues (W) are conserved (Fig. 1).

The WebLogo repeated sequence identification is based on the comparison of full-length PsMYB proteins. The height of the stack represents the conservation of the sequence at that position, and the height of a single letter at each position represents the relative frequency of the amino acid represented by that letter in the PsMYB protein sequence. The tryptophan residue (W) that represents the conservation of R2R3 is specifically labeled with a yellow pentagram whereas the first amino acid position of the easily replaceable R3 sequence is labeled with a colorless pentagram. To further investigate the conservation and diversity of, the 116 PsMYB TFs, their amino acid sequences were compared and an evolutionary tree was constructed using the functional grouping in *A. thaliana* as the reference for subgroup classification (Fig. 2A). The conserved motifs of the PsMYBs were analyzed using the Multiple Em for Motif Elicitation (MEME) website, and their motif compositions were predicted (Fig. S1). The visual results, showed that motif1, motif2, motif3, and motif4 were present in most PsMYB members. Motif1 and motif4 were conserved tryptophan residues in the R3 domain, whereas motif2 and motif3 were conserved tryptophan residues in the R2 domain. Motif1 and motif3 were present in all PsMYBs, further demonstrating the certainty and conservation of the *R2R3-MYB* gene family only in the P3 subgroup. The P24–P28 subgroups exhibited alternating repetitions of motif1 and motif3, the P27 subgroup exhibited three alternating repetitions, and the remaining subgroups exhibited two alternating repetitions. This indicates that they are quite conserved in PsMYBs. Furthermore, the exon sequences of the genes coding for five PsMYB proteins located in the P24 branch were not truncated by introns, whereas the exons and introns of the other genes alternated, and the P27 and P30 subgroups had the highest number of exons (Fig. 2B). Based on the analysis of conserved domains, all 116 PsMYBs possessed the PLN03091, PLN03212, or REB1 superfamily domains, which are significant features of MYB proteins. Among the 116 PsMYBs, 74 contained the PLN03091 superfamily, 46 contained the PLN03212 superfamily, and the PsMYB99 protein contained the REB1 and Alpah-mann_mid superfamilier. Additionally, the four genes of the P27 subgroup contained specific myb_DNA-bind_6 and myb_DNA-binding conserved domains. PsMYB7 contained not only the PLN03212 superfamily, but also an additional COG4913 superfamily (Fig. 2C). Intron–exon structure analysis indicated that the number of exons in the 116 *PsMYBs* ranged from 1 to 11, with genes with three exons being the most common. The structures of the same subgroup were highly similar (Fig. 2D).

Phylogenetic analysis and classification of the *R2R3-MYB* gene family

Phylogenetic analysis of the 116 PsMYB proteins in *P. sibirica* was performed using 126 MYB proteins of *A. thaliana* (*AtMYB*) downloaded from the *Arabidopsis* Information Resource for construction of the phylogenetic tree. The tree was divided into 30 subgroups (designated P1–P30) on the basis of the well-defined subgroups of the *R2R3-MYB* gene family in *A. thaliana* [12], and the specific conserved structure of the *P. sibirica* (Fig. 3).

Chromosomal location and synteny and Ka/Ks analysesof the *PsMYBs*

Detailed information of the chromosome position of each *PsMYB* genes was obtained on the basis of the *P. sibirica* genome annotation file, and the results were visualized. Overall, the 116 *PsMYBs* were unevenly distributed over eight chromosomes, with Chr1 being the longest. Chr3 contained most of the genes (total of 23). According to the chromosomal gene density analysis, the *PsMYBs* were mostly concentrated in areas with a high density of genes (Fig. 4).

In total, 28 pairs of duplicated *PsMYBs* were identified using the MCScanX program, identifying duplication events in genomes. Two pairs (*PsMYB67* and *PsMYB63*,

Table 1 Information on physicochemical properties of the PsMYBs and prediction of their subcellular localization

Gene Name	Gene ID	aa	MW (kDa)	pI	II	AI	GRAVY	Subcellular location
PsMYB1	PaF106G0100000483	254	28,642.96	5.07	51.63	64.17	-0.67	nucleus
PsMYB2	PaF106G0100000609	269	31,234.97	9.00	65.26	68.55	-0.89	nucleus
PsMYB3	PaF106G0100000611	254	29,610.23	9.06	57.84	62.20	-0.96	nucleus
PsMYB4	PaF106G0100001312	342	38,276.10	6.10	46.72	63.10	-0.72	nucleus
PsMYB5	PaF106G0100001431	241	27,479.25	9.49	57.63	80.21	-0.68	chloroplast
PsMYB6	PaF106G0100003058	219	25,008.12	6.53	47.62	75.25	-0.74	nucleus
PsMYB7	PaF106G0100003249	495	55,774.24	5.69	53.68	71.49	-0.67	nucleus
PsMYB8	PaF106G0100003312	464	51,178.39	6.11	62.01	52.35	-0.75	nucleus
PsMYB9	PaF106G0100003331	376	41,529.77	6.76	53.48	75.24	-0.51	nucleus
PsMYB10	PaF106G0100003723	305	34,450.66	6.45	66.89	68.49	-0.60	nucleus
PsMYB11	PaF106G0100003838	233	26,365.93	8.93	49.11	75.41	-0.83	nucleus
PsMYB12	PaF106G0100004751	202	23,644.73	8.90	58.18	72.92	-0.86	nucleus
PsMYB13	PaF106G0100004757	278	31,566.39	5.35	43.30	71.98	-0.71	nucleus
PsMYB14	PaF106G0100004758	284	32,237.32	9.05	42.03	65.32	-0.87	nucleus
PsMYB15	PaF106G0100004759	210	24,051.99	8.95	62.77	63.19	-0.96	nucleus
PsMYB16	PaF106G0100004885	292	31,911.66	6.80	43.81	59.83	-0.53	nucleus
PsMYB17	PaF106G0100004894	393	44,107.59	6.56	56.19	71.45	-0.67	nucleus
PsMYB18	PaF106G0100005026	312	33,823.03	8.65	55.01	69.71	-0.51	nucleus
PsMYB19	PaF106G0100005152	286	31,496.29	9.61	46.89	71.92	-0.54	chloroplast
PsMYB20	PaF106G0100005254	1038	113,836.23	5.26	56.64	63.13	-0.65	nucleus
PsMYB21	PaF106G0100005971	351	39,152.54	5.45	51.10	69.80	-0.67	nucleus
PsMYB22	PaF106G0100006366	226	25,962.25	9.82	54.49	70.44	-0.94	nucleus
PsMYB23	PaF106G0200006952	335	37,689.48	6.08	55.86	60.60	-0.66	nucleus
PsMYB24	PaF106G0200007454	560	60,746.45	5.11	50.07	65.52	-0.57	nucleus
PsMYB25	PaF106G0200008394	312	35,124.94	5.40	63.58	77.82	-0.60	nucleus
PsMYB26	PaF106G0200008877	595	66,283.66	9.19	60.58	65.90	-0.72	nucleus
PsMYB27	PaF106G0200008928	375	41,861.55	5.11	52.69	66.64	-0.69	nucleus
PsMYB28	PaF106G0200009263	295	33,548.76	9.26	42.73	68.71	-0.94	nucleus
PsMYB29	PaF106G0200009903	252	27,985.56	8.51	47.09	74.37	-0.64	nucleus
PsMYB30	PaF106G0200009999	243	27,020.37	6.10	57.98	75.84	-0.64	nucleus
PsMYB31	PaF106G0200010252	282	32,472.97	8.74	67.11	61.88	-0.82	nucleus
PsMYB32	PaF106G0200010338	245	27,680.29	5.33	50.11	76.08	-0.68	nucleus
PsMYB33	PaF106G0200010480	297	33,501.33	7.73	52.92	59.46	-0.80	nucleus
PsMYB34	PaF106G0200010860	431	47,519.79	6.57	46.00	73.60	-0.58	nucleus
PsMYB35	PaF106G0300011211	534	59,860.47	6.57	48.71	60.47	-0.80	nucleus
PsMYB36	PaF106G0300011220	198	22,637.41	6.66	48.38	72.98	-0.78	nucleus
PsMYB37	PaF106G0300011351	370	40,592.19	6.10	48.47	68.57	-0.47	nucleus
PsMYB38	PaF106G0300011470	396	43,306.92	6.10	56.08	66.74	-0.70	nucleus
PsMYB39	PaF106G0300011659	283	32,479.35	7.62	57.81	61.77	-0.91	nucleus
PsMYB40	PaF106G0300011882	369	41,744.74	4.97	54.96	76.67	-0.61	nucleus
PsMYB41	PaF106G0300011889	376	43,777.97	9.00	64.72	55.27	-1.01	nucleus
PsMYB42	PaF106G0300012032	376	41,269.76	9.51	49.04	57.98	-0.74	nucleus
PsMYB43	PaF106G0300012660	248	28,678.17	6.91	38.55	67.62	-0.83	nucleus
PsMYB44	PaF106G0300012662	181	21,062.21	9.91	48.16	76.46	-0.86	nucleus
PsMYB45	PaF106G0300012663	437	51,811.11	8.98	43.08	74.07	-0.57	nucleus
PsMYB46	PaF106G0300012665	243	28,345.89	6.91	48.55	66.21	-0.86	nucleus
PsMYB47	PaF106G0300012670	246	28,553.23	9.04	48.09	68.58	-0.83	nucleus
PsMYB48	PaF106G0300012674	244	27,468.97	6.76	51.08	73.57	-0.62	nucleus
PsMYB49	PaF106G0300012734	238	26,790.47	9.28	50.88	75.50	-0.73	nucleus
PsMYB50	PaF106G0300013636	328	36,725.28	8.09	44.14	70.55	-0.59	nucleus
PsMYB51	PaF106G0300013639	302	33,923.18	9.13	41.07	67.88	-0.63	nucleus
PsMYB52	PaF106G0300013643	328	36,766.52	8.83	41.11	70.27	-0.53	nucleus
PsMYB53	PaF106G0300014076	253	28,716.08	7.28	56.42	69.80	-0.70	nucleus

Table 1 (continued)

Gene Name	Gene ID	aa	MW (kDa)	pI	II	AI	GRAVY	Subcellular location
PsMYB54	PaF106G0300014123	290	32,928.89	5.66	42.35	70.24	-0.78	nucleus
PsMYB55	PaF106G0300014126	222	25,583.91	8.95	61.82	73.87	-0.88	nucleus
PsMYB56	PaF106G0300014434	270	30,377.98	4.94	62.41	73.37	-0.69	nucleus
PsMYB57	PaF106G0300014553	298	34,099.39	5.04	59.23	65.77	-0.86	nucleus
PsMYB58	PaF106G0400015268	336	37,121.20	5.20	57.33	66.49	-0.65	nucleus
PsMYB59	PaF106G0400015690	544	60,114.76	6.37	59.08	58.66	-0.66	nucleus
PsMYB60	PaF106G0400015691	528	57,457.76	7.29	65.62	64.15	-0.60	nucleus
PsMYB61	PaF106G0400015800	194	22,132.91	9.10	52.44	69.38	-0.99	nucleus
PsMYB62	PaF106G0400015803	224	25,540.70	8.60	51.78	79.20	-0.84	nucleus
PsMYB63	PaF106G0400015920	349	39,272.30	8.90	55.28	76.56	-0.58	nucleus
PsMYB64	PaF106G0400016645	365	41,153.37	4.84	55.11	60.96	-0.84	nucleus
PsMYB65	PaF106G0400016792	339	38,021.69	7.21	60.64	64.75	-0.70	nucleus
PsMYB66	PaF106G0400016875	329	37,139.67	6.14	49.15	68.75	-0.61	nucleus
PsMYB67	PaF106G0400017955	368	42,177.62	5.95	49.77	65.46	-0.84	nucleus
PsMYB68	PaF106G0400017967	251	28,447.92	6.19	57.73	66.53	-0.70	nucleus
PsMYB69	PaF106G0400018045	245	27,680.29	5.33	50.11	76.08	-0.68	nucleus
PsMYB70	PaF106G0400018116	202	23,192.47	9.44	56.87	66.09	-0.79	nucleus
PsMYB71	PaF106G0500018882	354	39,832.73	7.16	52.06	64.27	-0.66	nucleus
PsMYB72	PaF106G0500019249	164	18,601.88	9.46	66.44	68.48	-1.09	nucleus
PsMYB73	PaF106G0500019606	433	48,127.20	9.08	61.73	66.30	-0.62	nucleus
PsMYB74	PaF106G0500019746	346	39,162.82	5.83	61.73	66.30	-0.62	nucleus
PsMYB75	PaF106G0500019747	324	36,931.59	8.03	62.09	66.23	-0.65	nucleus
PsMYB76	PaF106G0500019998	285	32,533.65	5.73	44.19	71.86	-0.71	nucleus
PsMYB77	PaF106G0500020477	393	44,300.31	5.91	51.31	61.58	-0.67	nucleus
PsMYB78	PaF106G0500020533	481	54,722.90	6.03	34.92	60.17	-0.84	nucleus
PsMYB79	PaF106G0500020551	233	26,410.64	8.98	50.25	64.85	-0.82	nucleus
PsMYB80	PaF106G0500020648	360	40,461.84	5.97	56.84	71.08	-0.52	nucleus
PsMYB81	PaF106G0500020762	363	40,464.92	6.53	46.11	66.17	-0.82	nucleus
PsMYB82	PaF106G0500020797	330	36,940.36	6.74	45.21	74.00	-0.64	nucleus
PsMYB83	PaF106G0500021122	316	35,624.04	6.24	75.59	69.81	-0.59	nucleus
PsMYB84	PaF106G0500021206	223	25,467.23	9.31	51.02	76.55	-0.79	nucleus
PsMYB85	PaF106G0500021207	148	17,141.57	9.98	30.19	64.59	-1.08	chloroplast
PsMYB86	PaF106G0500021313	261	30,170.71	7.74	75.36	64.33	-0.84	nucleus
PsMYB87	PaF106G0600021449	400	42,602.06	6.97	46.37	68.67	-0.59	chloroplast outer membrane
PsMYB88	PaF106G0600021756	325	36,137.54	6.83	40.22	62.18	-0.75	nucleus
PsMYB89	PaF106G0600022076	418	46,615.09	7.78	52.74	57.20	-0.99	nucleus
PsMYB90	PaF106G0600022375	261	29,714.31	6.91	66.25	59.08	-0.85	nucleus
PsMYB91	PaF106G0600022680	365	41,128.66	7.24	51.47	55.84	-0.86	nucleus
PsMYB92	PaF106G0600023626	244	28,368.13	7.66	62.68	77.46	-0.74	nucleus
PsMYB93	PaF106G0600023629	242	28,077.70	9.47	46.80	69.30	-0.83	nucleus
PsMYB94	PaF106G0600023630	243	28,240.81	9.51	43.06	71.03	-0.84	nucleus
PsMYB95	PaF106G0600023821	343	38,111.72	5.93	51.96	70.32	-0.62	nucleus
PsMYB96	PaF106G0600024116	377	41,992.74	7.15	50.82	67.32	-0.63	nucleus
PsMYB97	PaF106G0600024275	331	35,708.30	8.98	65.92	73.60	-0.47	nucleus
PsMYB98	PaF106G0600024570	593	64,869.37	7.13	58.92	61.21	-0.72	nucleus
PsMYB99	PaF106G0600024579	356	41,053.82	9.34	62.41	73.46	-0.89	nucleus
PsMYB100	PaF106G0600024701	395	43,293.08	5.23	38.34	64.99	-0.60	nucleus
PsMYB101	PaF106G0600025148	463	51,515.75	7.63	56.86	57.95	-0.78	nucleus
PsMYB102	PaF106G0600025755	266	30,366.97	5.70	54.38	66.02	-0.91	nucleus
PsMYB103	PaF106G0700026197	291	32,775.67	8.43	63.99	64.81	-0.73	nucleus
PsMYB104	PaF106G0700026552	262	29,128.53	6.44	52.15	63.32	-0.59	nucleus
PsMYB105	PaF106G0700026725	256	29,416.84	5.73	46.03	73.12	-0.81	nucleus
PsMYB106	PaF106G0700026726	223	25,155.26	9.46	44.60	63.05	-0.87	nucleus

Table 1 (continued)

Gene Name	Gene ID	aa	MW (kDa)	pI	II	AI	GRAVY	Subcellular location
PsMYB107	PaF106G0700026727	323	36,028.34	6.94	49.94	70.37	-0.78	nucleus
PsMYB108	PaF106G0700027598	321	36,171.50	8.62	58.87	65.39	-0.71	nucleus
PsMYB109	PaF106G0700027623	248	28,867.04	5.43	53.63	62.18	-0.91	nucleus
PsMYB110	PaF106G0700029014	348	39,559.43	7.11	54.25	73.99	-0.78	nucleus
PsMYB111	PaF106G0700029057	414	45,517.58	8.23	47.81	57.51	-0.75	nucleus
PsMYB112	PaF106G0800029300	545	60,604.42	4.98	47.88	63.63	-0.79	nucleus
PsMYB113	PaF106G0800029883	264	29,573.93	4.99	51.30	62.88	-0.70	nucleus
PsMYB114	PaF106G0800030628	256	28,337.17	8.84	47.07	77.03	-0.61	nucleus
PsMYB115	PaF106G0800030962	302	32,918.56	6.38	50.39	66.59	-0.67	nucleus
PsMYB116	PaF106G0800032722	194	22,563.99	9.56	61.31	50.82	-1.12	nucleus

and *PsMYB68* and *PsMYB63*) were tandem replicates, whereas the rest were segmental duplicates (Fig. 5). During the evolution of the *P. sibirica* genome, the number of segmental duplication events in the *PsMYB* gene family was much greater than the number of tandem replication events, with only the two pairs mentioned above belonging to the tandem replication group. The synonymous substitution rate (Ks), non-synonymous substitution rate (Ka), and Ka/Ks ratio were calculated to determine the selection pressure during the evolution of the *PsMYB* family after gene duplication events. Except for the duplicate gene pairs without numerical values, the remaining 20 duplicate gene pairs all had Ka/Ks ratios below 1 (between 0.11 and 0.66), indicating that these gene pairs have been purified and screened (Table 2).

The evolutionary relationships of the *MYB* gene family members of *P. sibirica* with those of *A. thaliana*, *P. mume*, *P. persica*, *P. salicina*, and *P. avium*, as well as their collinearity, were analyzed. According to the results shown in the graph, the number of collinear pairs between the genome of *P. sibirica* and those of other *Rosaceae* plants was greater than that between the *P. sibirica* and *A. thaliana* genomes. The genetic relationship between *P. sibirica* and *P. mume* was the closest (Fig. 6).

Cis-acting elements analysis of *PsMYBs*

To better understand the specific types and distribution of *cis*-acting elements present in the *PsMYBs* promoter, *cis*-acting elements were predicted on the basis of the 2000 bp DNA sequence upstream of the *R2R3-MYB* coding sequence extracted from the PlantCARE database. *Cis*-acting elements can be divided into three categories according to their regulatory functions: plant growth/development, plant hormone responses, and biotic/abiotic stress responses. This study prioritized screening abiotic stress-related *cis*-acting elements, such as those related to hormone responses and low-temperature stress (Table S1). Furthermore, at least one abiotic stress-responsive *cis*-element was found in each *PsMYB* promoter region, with methyl jasmonate (MeJA) and abscisic acid (ABA) being the most widely used *cis*-responsive elements. This further demonstrated that members of the *R2R3-MYB* subgroup primarily regulate abiotic stress responses in plants. Furthermore, 95 *PsMYBs* had 331 elements involved in the ABA response, 87 *PsMYBs* had 314 elements related to MeJA, and 65 *PsMYBs* had 93 elements related to gibberellins. In addition to hormone responses, 61 *PsMYBs* were involved in drought induction. Among the 37 *PsMYBs* involved in cold related responses, *PsMYB41* contained the most low-temperature response (LTR) elements, whereas *PsMYB29*, *PsMYB54*, *PsMYB61*, *PsMYB87*, *PsMYB98*, and *PsMYB102* contain two low temperature response (LTR) elements. *PsMYB62* contained22 elements, making it the gene with the highest number of *cis*-acting elements, 17 of which were related to hormone responses. *PsMYB104* contained up to nine *cis*-acting elements (Fig. 7).

Gene Ontology annotation analysis of PsMYBs

The EggNog website for Gene Ontology (GO) annotation analysis was used to investigate the functions of 75 of the 116 PsMYB TFs. Under the organelle terms of the Cellular Components category, the 75 PsMYBs were annotated to the nucleus. Their molecular functions included transcriptional regulation and binding. The 37 PsMYBs showed responded to exogenous stimuli (GO: 0009410), while 25 responded to jasmonic acid (GO: 0009753), and 20 responded to abscisic acid (GO: 0009737), respectively. Six genes (PsMYB1, PsMYB41, PsMYB68, PsMYB81, PsMYB98, and PsMYB113) were annotated as responsive to cold (GO: 0009409) (Table S2) (Fig. 8).

Protein interactions of PsMYBs

Using the online tool STRING, we predicted protein interaction networks on the basis of the homologous relationships between PsMYBs and AtMYBs. The interactions are displayed using the gene names of *A. thaliana* (Fig. 9), and detailed information is listed in Table S3. In total, 82 PsMYBs interacted with more than five AtMYBs, with AtMYB1 (PsMYB100) having the most interactive relationships (30 in total).

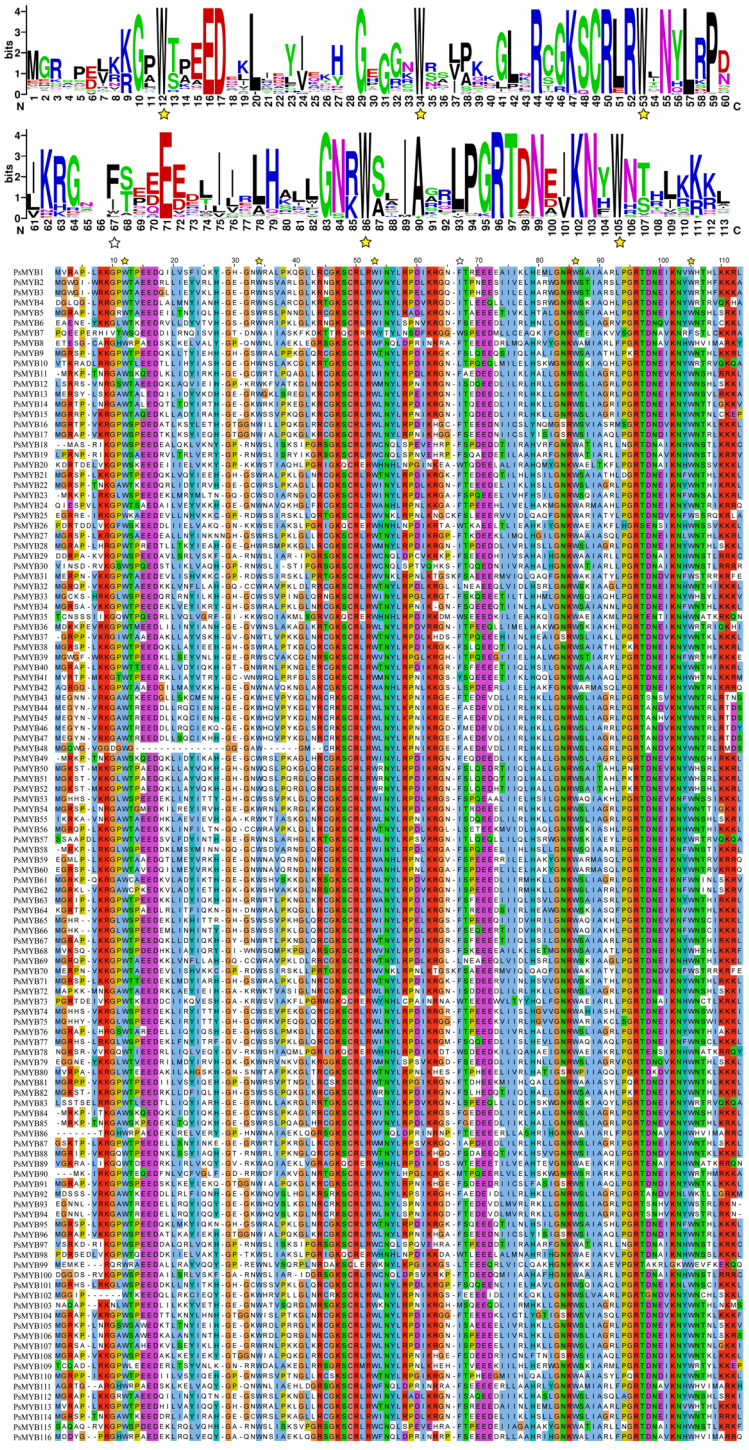

Fig. 1 Multiple sequence alignment and WebLogo diagram of the R2R3-MYB conserved domain in PsMYB transcription factor. Multiple sequence comparison refers to comparing the conserved domains of all PsMYB proteins

Transcriptome analysis of *PsMYBs*

We compared downloaded data of the pistil transcriptomes of the cold-tolerant "NO.453" and cold-sensitive "NO.371" clone groups, observing the differential expression of their *PsMYB* genes without stress and their changes in differential expression with increasing stress time (Table S4). As shown in the heatmap (Fig. 10), the expression levels of *PsMYB1*, *PsMYB9*, *PsMYB82* and *PsMYB91* gradually increased with increasing cold stress duration, indicating a progressive phenomenon. *PsMYB43*, *PsMYB48*, and *PsMYB62* were already highly expressed before the clone groups were subjected to cold

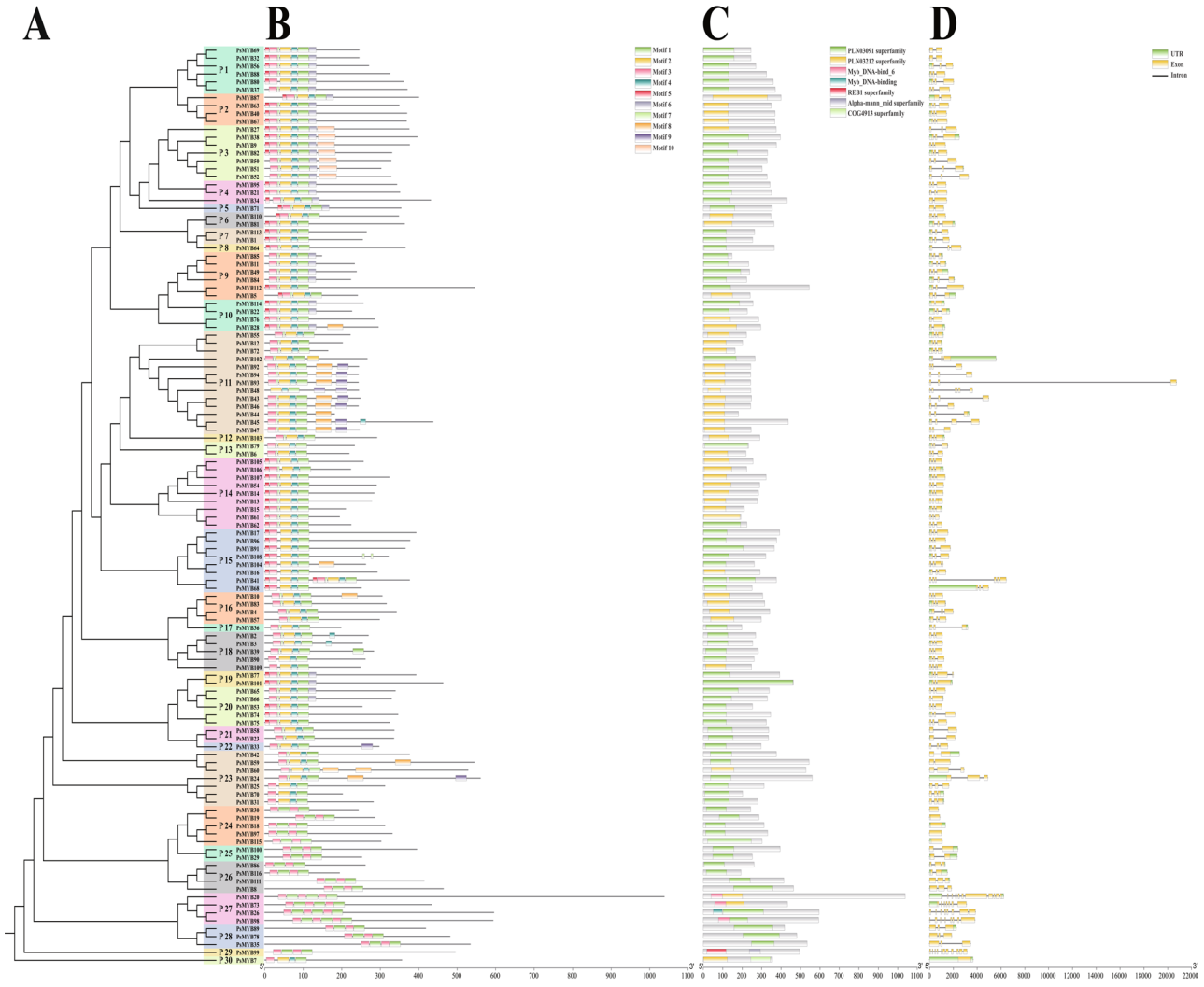

Fig. 2 Phylogenetic relationships **(A)**, conserved motifs **(B)**, conserved domains **(C)**, and gene structure analysis **(D)** of *PsMYBs*

stress and showed a significant decrease in expression under stress. When neither clone group was subjected to cold stress (0 h), the expression levels of the various genes in "NO.453" were either higher or lower than those of the genes in "NO.371". With increasing time under cold stress, the expression levels of *PsMYB40* and *PsMYB61* were remarkably higher at 15 min than at 0 h. The gene expression levels of *PsMYB44* and *PsMYB50* remarkably increased at 1 h, while those of *PsMYB6* and *PsMYB8* were remarkably increased at 2 h. By contrast, the expression levels of *PsMYB10*, *PsMYB69*, and *PsMYB86* were remarkably decreased with increasing stress time.

Expression analysis of *PsMYBs* in different tissues

To characterize the potential functions of *PsMYBs*, we screened 22 *PsMYBs* that may be related to plant growth and development using promoter *cis*-acting elements, GO annotation analysis, and the transcriptomes, and other homologous *MYB* genes from plants (Table S5) and analyzed their expression characteristics using qRT-PCR. The total *18S rRNA* from all experimental materials was extracted as the internal standard. After normalizing all experimental expression levels against the reference *18S rRNA* gene, the gene expression levels in the roots were used as the control groups to compare differences to the gene expression levels in the other plant parts and roots. The 22 *PsMYBs* exhibited different expression patterns in different tissues. Based on the heatmap and bar chart results, four genes (*PsMYB8*, *PsMYB82*, *PsMYB98*, and *PsMYB99*) were significantly overexpressed in the pistil, with *PsMYB99* being the most significantly, affected (upregulated 25 times more compared with the control level), followed by *PsMYB82* (upregulated 4.6 times more compared with the control level). Five genes (*PsMYB19*, *PsMYB22*, *PsMYB38*, *PsMYB41*, *PsMYB45*, and *PsMYB99*) were significantly overexpressed in the stem, with *PsMYB99* showing a 5-fold higher level of expression relative to the control level. *PsMYB19* and

Fig. 3 Phylogenetic analysis of *P. sibirica* and *A. thaliana* MYB domains. The bootstrap was set to 1000. Different coloured branches represent different subgroups. *P. sibirica* and *A. thaliana* are marked as black pentagrams and purple circles, respectively

PsMYB41 were significantly expressed in the leaves, whereas *PsMYB22* and *PsMYB44* were highly expressed in the petals (Fig. 11).

Expression analysis of *PsMYBs* sunder low-temperature stress

The qRT-PCR was used to detect the expression levels of 22 *PsMYBs* at different stages of low-temperature stress and investigate the expression patterns of *PsMYBs* that may be related to the cold response. The cDNA plant materials with different cold stress times are all pistils. After normalizing all expression data against the internal control, gene expression at 0 h was used as the control group to observe expression differences at the other cold stress times. According to the heat map (Fig. 12), the 22 *PsMYB* genes showed significant differences in expression, with the majority exhibiting high expression levels after low-temperature stress. As shown in the bar chart, *PsMYB99* expression rapidly increased by 55 times after 15 min of cold stress treatment, *PsMYB57* expression increased by 30 times after 1 h of cold stress treatment, and *PsMYB19*, *PsMYB22*, *PsMYB67*, and *PsMYB108* expression showed a 10-fold increase after 1–2 h of cold stress treatment relative to the 0 h levels (Fig. 12). The

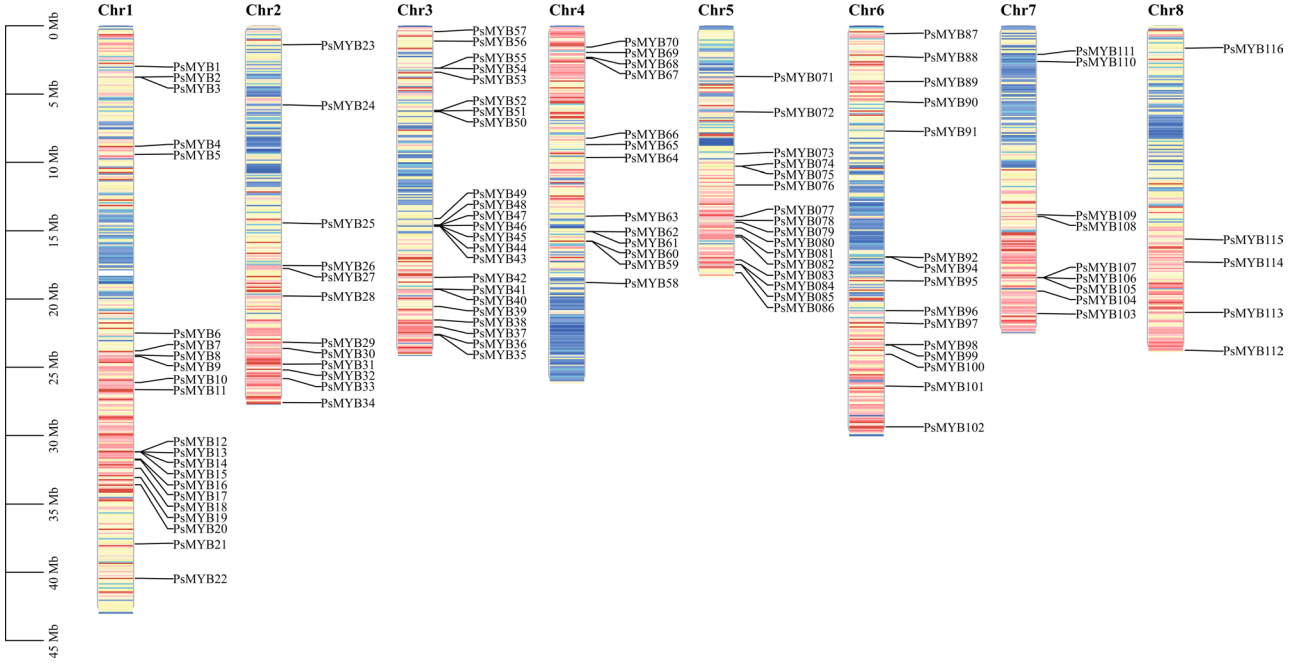

Fig. 4 Chromosomal locations of the *PsMYB* genes. Chr1–Chr8 represent different chromosomes. The colors on the chromosomes indicate levels of gene density, with red representing high and blue representing low gene density in that region

results indicate that *PsMYB99* and *PsMYB57* are genes that respond rapidly to cold stress.

Discussion

TFs are key critical components of the inducible signal transduction networks [27]. They are widely expressed in organisms, with activators and repressors displaying modular structures [12]. They play essential roles in the regulation of plant growth and development and responses to biotic and abiotic stress [28, 29]. Transcription factors typically consist of DNA-binding domains, nuclear localization signals, transactivation domains, and oligomeric sites [30, 31]. They are divided into different gene families, including *AP2/ERF*, *EIN3*, *MYB*, *bZIP*, *bHLH*, *WRKY*, and *NAC* according to their different DNA-binding domains. MYB is one of the largest and most abundant of the TF families in plants [1].

Systematic analyses of the *MYB* gene family at the whole-genome level have been reported for various plants [32]. *R2R3-MYB* is the most common subgroup of in the *MYB* gene family [10], with members of the 2R subfamily mainly controlling plant growth, development, and stress response [22]. This gene family has been identified and studied in many higher plants in the *Rosaceae* family, with numbers ranging from tens (96 found in *P. salicina*) [33] to over 200 (219 found in *Malus domestica*) [8]. In this study, we identified 116 *PsMYBs*, which were similar to statistical data on the *R2R3-MYB* gene subfamily in other *Rosaceae* tree species [8], indicating the highly conserved nature of this gene family in the *Rosaceae* family.

Comparisons of the physicochemical properties of the various PsMYB proteins revealed significant differences in their numbers of amino acids, relative molecular weights, and isoelectric point values. The acidity and alkalinity of the different TFs also differed, further showing the significant differences in their physical properties. The GRAVY values of the 116 PsMYBs were all negative, indicating they were hydrophilic proteins. Hydrophilicity can prevent harmful changes in the secondary and tertiary structures of proteins and protects cellular components from the adverse effects caused by low water utilization rates (e.g., water present under freezing conditions) [34]. PsMYB20 had the highest number of amino acids (1038 in total) and relative molecular weight (113.8 kDa), which were much higher than the average number of amino acids (319.1) and relative molecular weight (35.9 kDa) of the other proteins. The TF probably forms more complex protein structures and participates in more plant physiological functions. Of all PsMYBs, 96.6% were predicted to be localized in the nucleus, indicating that they are mainly involved in transcriptional regulation within this organelle. The remaining four PsMYBs, located in the outer membrane of the chloroplast, are speculated to be involved in gene transcription and expression related to chloroplast formation or photosynthesis.

Multiple sequence alignments and WebLogo indicated that *P. sibirica* contains a highly conserved R2R3

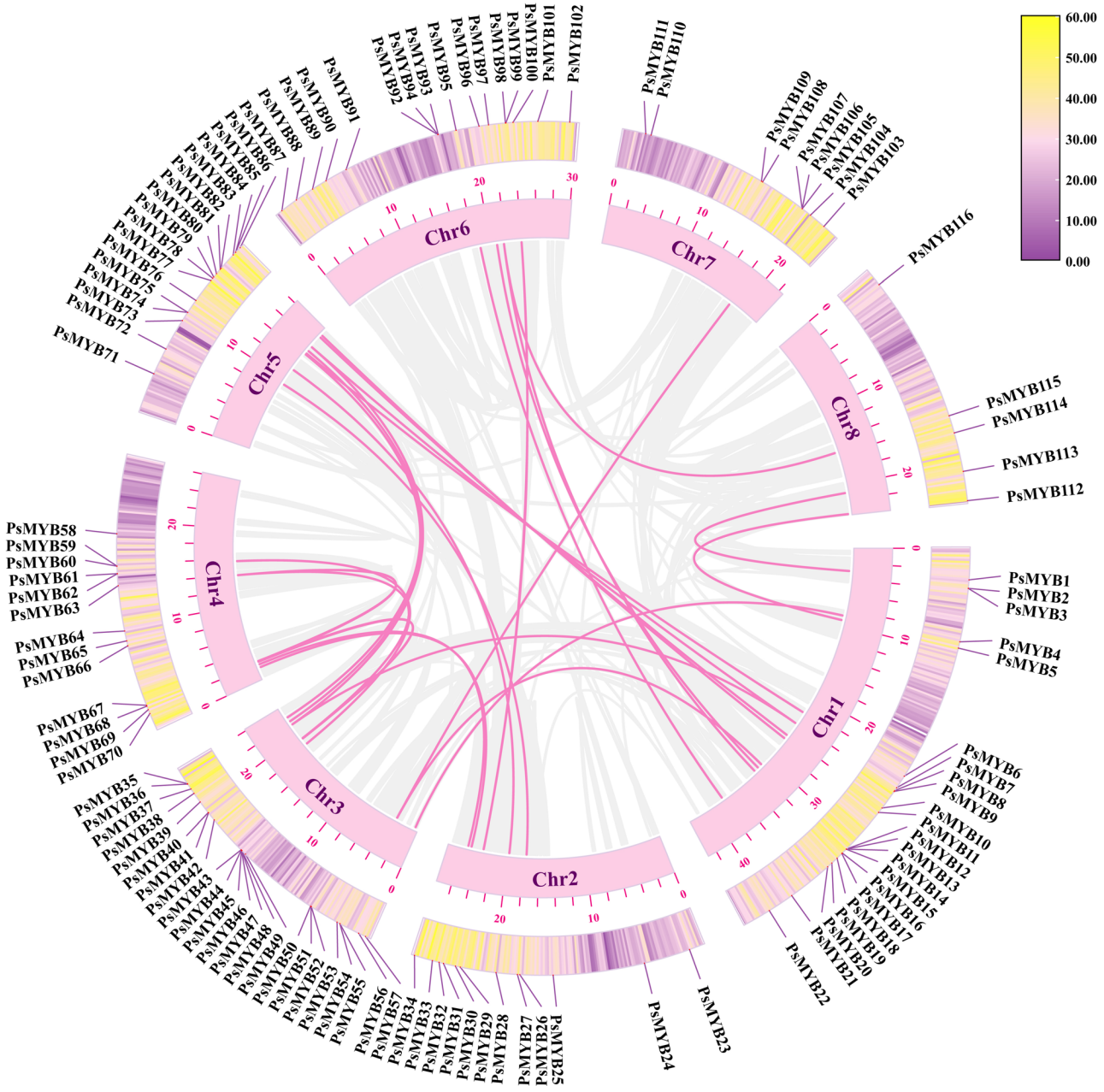

Fig. 5 Chromosomal distribution of the *PsMYB* genes and their repetitive events. The 8 chromosomes of *P. sibirica* are denoted as chr1-chr8. All identified duplicate segments are represented in gray, with the pink line indicating *PsMYB* duplication segments

domain with three highly conserved tryptophan residues in the R2 region. The first tryptophan residue in the R3 region was replaced with phenylalanine, isoleucine, and leucine, similar to the R3 structure of other species [35, 36]. PsMYB41 has two duplicates of motif1 and motif5, making it the only gene with such a duplication. The *PsMYB41* gene may have undergone mutations or duplication in a conserved domain sequence during evolution. According to the prediction of conserved domains in the PsMYB proteins, all were shown to contain domains from the PLN03091, PLN03212, or REB1 superfamilies, which represent significant *R2R3-MYB* conserved domain features, once again proving the accuracy of gene screening [37]. The number of exons in *PsMYBs* ranged from 1 to 11, with the most gene members having three exons, which is consistent with previous reports [11, 24].

The phylogenetic tree constructed on the basis of *P. sibirica* motif functions, *A. thaliana* classification, and the 116 *PsMYBs* comprised 30 subgroups. The TFs encoded by genes in the same subgroup may have similar structures and functions, that with those are more similar during branching also tending to be more similar

Table 2 Ka/Ks ratios for the duplicate pairs of *PsMYBs*

Gene Name	Gene Name	Ka	Ks	Ka_Ks	Type of duplication
PsMYB12	*PsMYB55*	0.28	NaN	NaN	Segmental
PsMYB13	*PsMYB54*	0.42	2.94	0.14	Segmental
PsMYB4	*PsMYB57*	0.33	2.81	0.12	Segmental
PsMYB9	*PsMYB38*	0.25	1.69	0.15	Segmental
PsMYB6	*PsMYB79*	0.35	1.47	0.24	Segmental
PsMYB10	*PsMYB83*	0.39	2.70	0.15	Segmental
PsMYB11	*PsMYB84*	0.33	1.51	0.22	Segmental
PsMYB9	*PsMYB82*	0.43	2.30	0.19	Segmental
PsMYB17	*PsMYB96*	0.31	1.43	0.21	Segmental
PsMYB18	*PsMYB97*	0.23	1.23	0.19	Segmental
PsMYB21	*PsMYB95*	0.27	1.24	0.22	Segmental
PsMYB1	*PsMYB113*	0.26	2.41	0.11	Segmental
PsMYB5	*PsMYB112*	0.31	2.19	0.14	Segmental
PsMYB31	*PsMYB70*	0.01	0.01	0.66	Segmental
PsMYB28	*PsMYB76*	0.33	1.68	0.20	Segmental
PsMYB29	*PsMYB100*	0.30	1.30	0.24	Segmental
PsMYB41	*PsMYB68*	0.45	NaN	NaN	Segmental
PsMYB40	*PsMYB67*	0.26	NaN	NaN	Segmental
PsMYB41	*PsMYB67*	0.57	NaN	NaN	Segmental
PsMYB42	*PsMYB60*	0.45	1.99	0.23	Segmental
PsMYB40	*PsMYB63*	0.36	2.75	0.13	Segmental
PsMYB35	*PsMYB78*	0.49	NaN	NaN	Segmental
PsMYB37	*PsMYB80*	0.40	1.93	0.21	Segmental
PsMYB38	*PsMYB82*	0.40	1.95	0.20	Segmental
PsMYB55	*PsMYB107*	0.44	NaN	NaN	Segmental
PsMYB67	*PsMYB63*	0.29	NaN	NaN	Tandem
PsMYB68	*PsMYB63*	0.57	NaN	NaN	Tandem
PsMYB97	*PsMYB115*	0.33	1.23	0.27	Segmental

[38]. In the P2 subgroup, the *AtMYB49*-encoded TF regulates keratin deposition and the antioxidative capacity in leaves to enhance salt-alkali resistance in the plant [39], whereas in the same subgroup, the *AtMYB41* responds to abiotic stressors such as cold, ABA, and salt treatment by controlling cell expansion and stratum corneum deposition [40, 41]. The *Vaccinium corymbosum* gene *VcMYB171* (an *AtMYB41* homologous gene) is predicted to respond to drought by altering leaf morphology [42]. P2 subgroup genes, such as *PsMYB40*, *PsMYB63*, *PsMYB67*, and *PsMYB87*, also control cellular keratin deposition to cope with external pressure. The *AtMYB31*, *AtMYB94*, and *AtMYB96* of the P6 subgroup have been found to be associated with the biosynthesis of epidermal wax in reproductive tissues, thereby participating in seed synthesis [43, 44] *ZmFDL1/MYB94* (*AtMYB94*-like gene)participates in wax biosynthesis to cope with water stress in *Z. mays* [45]. These findings support the prediction predicting that P6 subgroup *PsMYBs* also respond to expression in reproductive cells. The *AtMYB15* in the P7 subgroup acts as a negative regulator of freezing tolerance by inhibiting the expression of *CBF1/DREB1* [46, 47], whereas PsMYB1 and PsMYB113 in the same subgroup may also be negative regulators of freezing tolerance. According to previous reports, AtMYB11, AtMYB12, and AtMYB111 in the P9 subgroup control the biosynthesis of flavonols [48]. Overexpression of the *MrMYB12* gene (an *AtMYB12*-like gene) from *Morella rubra* in tobacco plants can increase the accumulation of flavonols in the transgenic line [31]. Therefore, PsMYB85, PsMYB11, PsMYB49, PsMYB84, PsMYB112, and PsMYB5, which are all in the same P9 subgroup, may be involved in flavonols synthesis.

Gene duplication is one of the most important mechanisms for emerging new gene functions, playing a crucial role in expanding plant gene families [49], and acting as a driving force for evolution [50]. During the evolution of the *R2R3-MYB* genome in *P. sibirica*, 28 pairs of genes were involved in duplication events, with 92.9% having undergone segmental duplication. In pears, 50 gene pairs were all found to have undergone segmental duplication [24]. Among the 37 gene pair duplication events in *Theobroma cacao*, 33 were of segmental type [51]. Segmental duplication events were also confirmed to play a crucial role in expanding the *SmR2R3-MYB* gene family in Eggplant Fruit (*Solanum melongena*) [52]. In summary, segmental duplication events may be the main driving force of evolution. Except for *PsMYB55* (subgroup P11) and *PsMYB107* (subgroup P14), all the other gene replicates fall under the same subgroup. After segmental duplication, the *PsMYB55* and *PsMYB107* may have functional differences owing to changes in their structure and motif composition, whereas the other TFs still have similar structures and motif compositions after the gene duplication, allowing them to remain consistent in function and have the same regulation and control.

Amino acid substitutions can be defined as non-synonymous (Ka) or synonymous (Ks) depending on whether the changes in the codon sites alter the translation of the amino acid [53]. The evolutionary selection type of genes can be determined on the basis of their Ka/Ks ratio, where a value less than 1 indicates that these genes are well conserved and will continue to evolve under purifying selection pressure, whereas a value greater than 1 indicates positive selection and a possible loss of the ancestral RNA editing site, resulting in new evolution [54]. In this study, the Ka/Ks ratios of the 20 duplicate gene pairs were all less than 1, indicating purifying selection. These well-preserved genes continue to evolve steadily and conservatively.

Homology analysis can intuitively reveal the collinearity of orthologous MYB compounds in *Rosaceae* plants [24]. The results of this study showed that the number of collinear gene pairs between *Rosaceae* plants and *P. sibirica* was much higher than that between *P. sibirica* and *A. thaliana*, which is consistent with the performance of *P. salicina* [33]. *R2R3-MYB* genes may have existed as ancient conserved genes before the

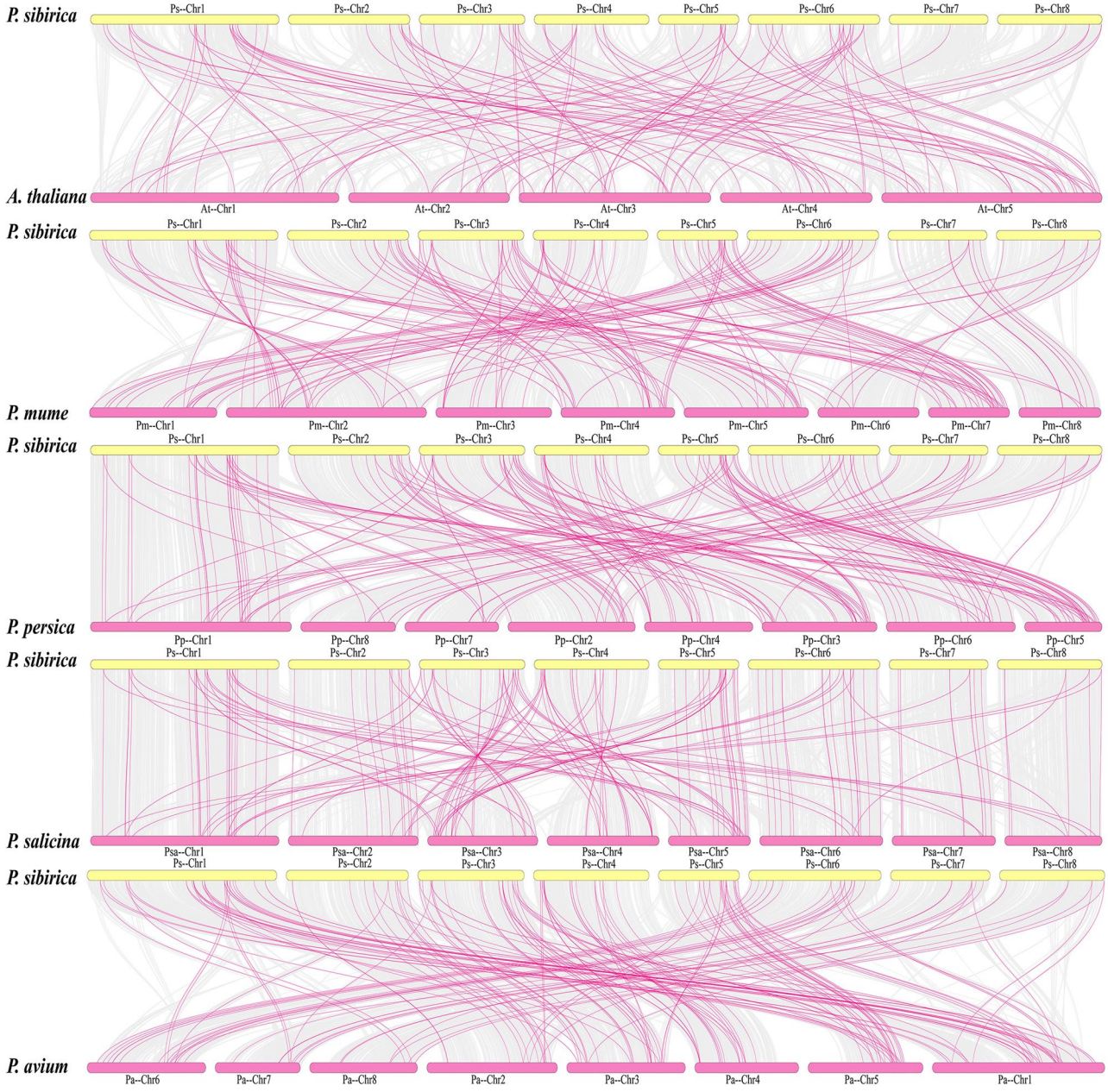

Fig. 6 *MYB* gene homology between *P. sibirica* and five representative plants. The plants named with the different prefixes "At", "Pm," "Pp," "Psa," and "Pa" represent *A. thaliana*, *P. mume*, *P. persica*, *P. salicina*, and *P. avium*, respectively. The gray line represents collinear gene pairs between the genomes of *P. sibirica* and those of the other plants, whereas the pink line represents highlighted *MYB* gene pairs

differentiation of the *Rosaceae* species. *P. sibirica* and *P. avium* shared the highest number of collinear pairs and exhibited the most similar genetic relationships. Further research on the evolutionary processes and patterns of the *Rosaceae* family can be conducted on the basis of the study of these gene pairs.

TFs receive internal and external signals in a sequence-specific manner, activating or inhibiting target gene transcription through the transcriptional regulation of gene expression upon binding to *cis*-regulatory elements in the gene promoter region [55]. The upstream promoter region of *PsMYBs* contains many potential regulatory *cis*-elements that respond to either MeJA or ABA, the latter of which has the highest number of such elements. ABA has been shown to be a key regulator in the response of plants to abiotic stress [56]. MeJA supplementation under cold stress conditions can reduce oxidative stress markers, and the jasmonate also acts as an inducer to enhance plant tolerance to cold stress [57, 58]. GO of the *PsMYBs* annotation also showed similar results, with most genes annotated as being related to MeJA and ABA. Furthermore, *cis*-acting elements involved in LTR were

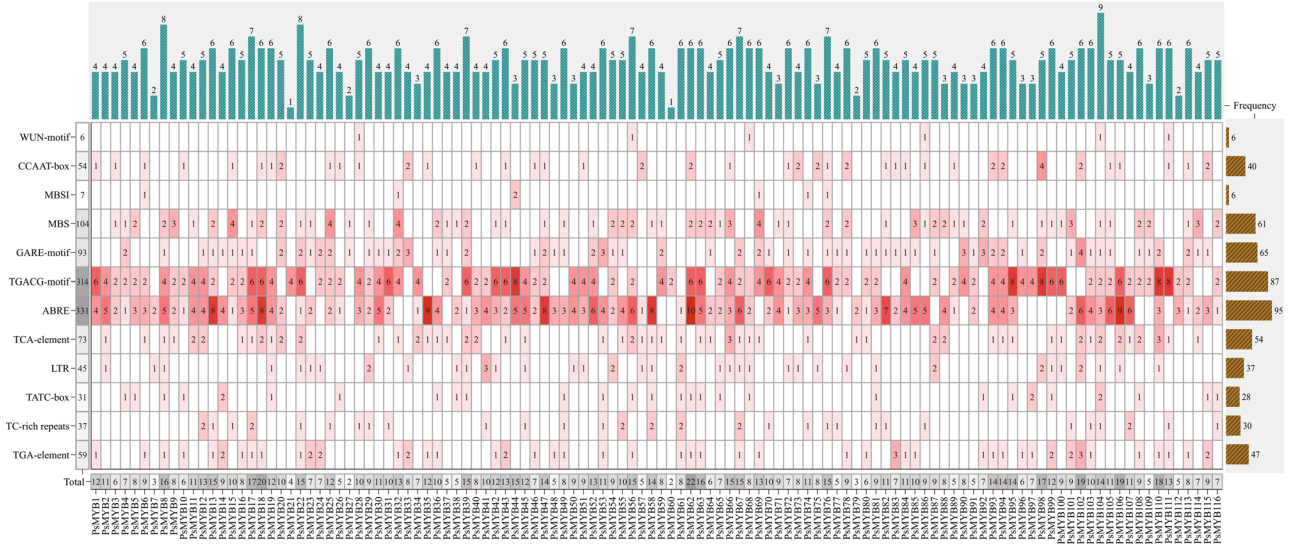

Fig. 7 Predict the *cis*-elements in the *PsMYB* gene promoter. The heatmap represents the number of *cis*-acting elements in a gene, whereas the bar chart represents the frequency of gene or element occurrence. The text in the square represents the total number of *cis*-acting elements for each gene

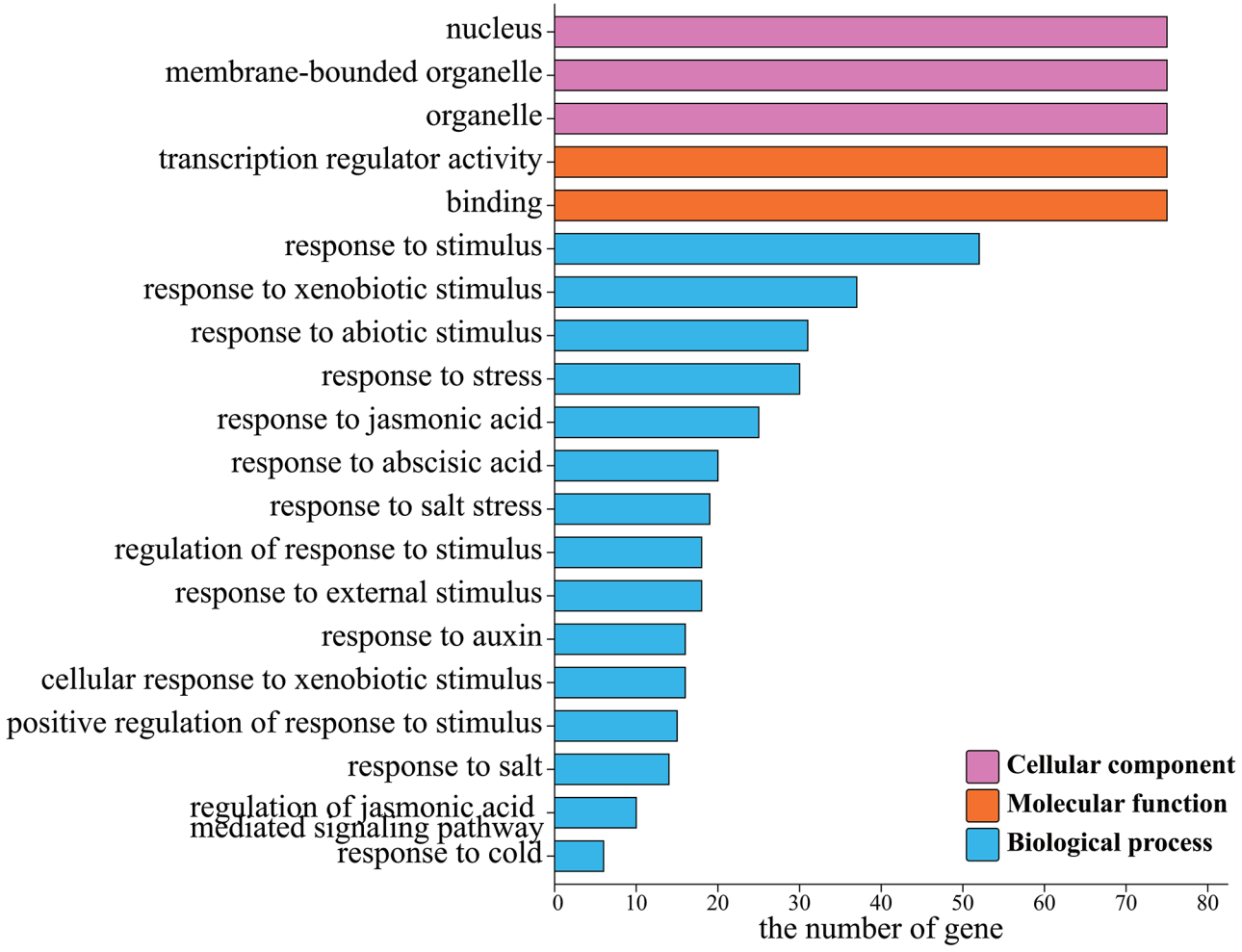

Fig. 8 Gene ontology (GO) annotation of the PsMYBs, showing enrichment in the Cellular component, Molecular function and Biological process categories

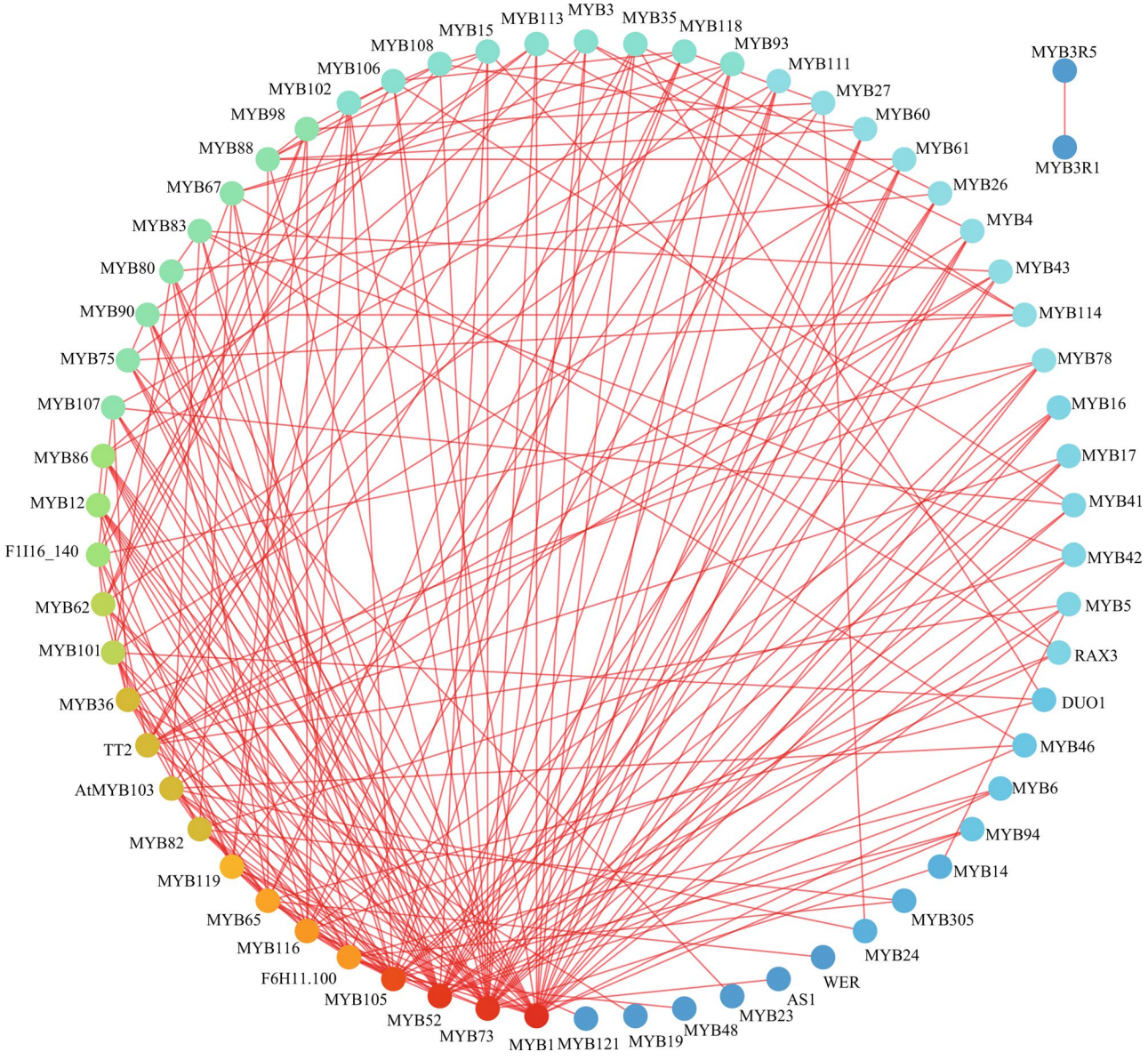

Fig. 9 Protein interaction network for the PsMYBs based on MYB orthologs in *A. thaliana*. The red line represents the interaction relationship, the circle color represents the number of interacting genes, with red denoting more and blue denoting less interactions

found in 37 *PsMYBs*, among which *PsMYB41* contains three LTR elements that may be significantly expressed in cold environments.

From the protein interaction prediction results, AtMYB1 was found to have the most interactive relationships with the PsMYB TFs in this study. According to previous reports, AtMYB1 negatively affects *A. thaliana* seed germination under saline-alkali conditions by regulating ABA levels [59]. PsMYB100, which interacted with AtMYB1, can also affect *P. sibirica* seed germination by regulating ABA levels. AtMYB94, which interacted with and belonged to the same subgroup as PsMYB81, enhances plant resistance to external stimuli by promoting epidermal wax [43]. GO analysis of PsMYB81 annotated it to terms to cold stress, indicating that the TF may also regulate the production of epidermal wax to resist external stress.

Plants exhibit differential expression of multiple regulatory genes under abiotic stress conditions, forming a complex regulatory network. Transcriptome sequencing can be used to quantify gene expression and identify functional genes [60]. The TF encoded by *AtMYB15* (a *PsMYB1* homologous gene) is a negative regulator of cold signaling [47]. In the comparison of transcriptomes following low temperature stress, *PsMYB1* showed a significant and gradual upregulation in expression, indicating that it belongs to the cold resistance genes. Transgenic *A. thaliana* plants containing the *FvMYB82* gene from

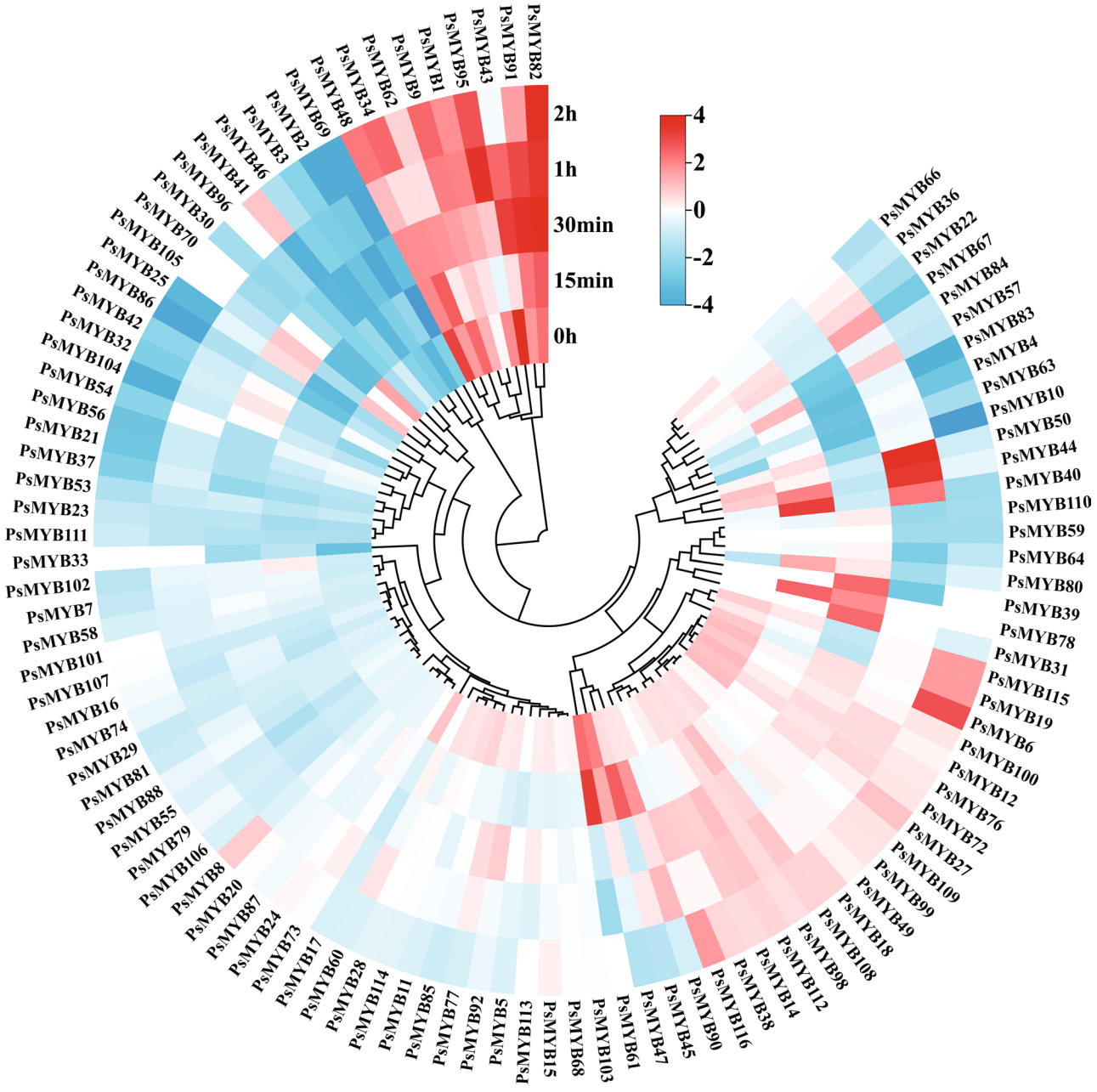

Fig. 10 Genes of cold-resistant "NO.453" and cold sensitive "NO.371" clones show differential expression at different durations of cold stress. The color represents the transcriptome difference between the two clones, with red indicating higher gene expression in "NO.453" than in "NO.371", and blue indicating the opposite trend

Fragaria vesca showed significantly improved tolerance to salt and cold stress [22]. However, the expression of *PsMYB103*, a homologous gene, was downregulated in the transcriptome with increasing stress time. The differences in these results suggest that the expression patterns of homologous genes may differ between herbaceous and woody plants. *A. thaliana AtMYB16* and *AtMYB106* (a *PsMYB9* homologous genes) regulate the stratum corneum formation in reproductive organs and trichomes, thereby protecting plants from external damage [61]. According to the transcriptome results, *PsMYB9* expression was significantly upregulated under cold stress, suggesting that the TF may regulate stratum corneum generation in plants and enhance their cold resistance under low temperature stress. The *EqMYB157* gene (a *PsMYB78* homologous gene) of oil palm (*Elaeis guineensis*) can improve the tolerance of *A. thaliana* to cold stress [62]. In our study, the expression of the *PsMYB78* in the transcriptome increased after 15 min of cold stress, suggesting that it may have a regulatory effect on plant

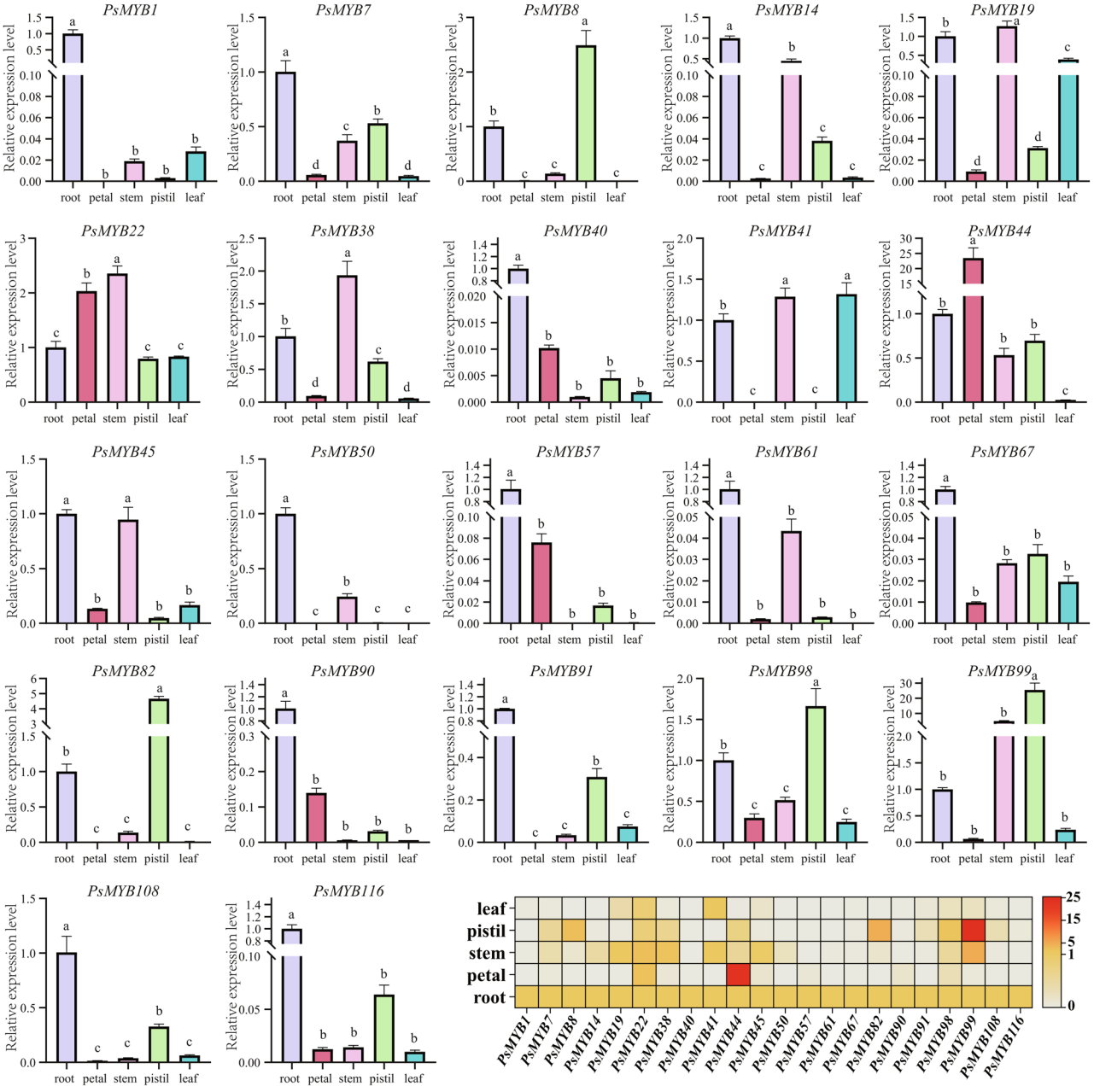

Fig. 11 Expression of 22 *PsMYBs* in different organs of *P. sibirica*. Results are the average of three replicates. The vertical bars indicate the standard deviation, and different letters over the columns denote significant differences ($p < 0.05$)

cold tolerance. The *AtMYB93* gene (a *PsMYB95* homologous gene) can regulate lateral root development [63], whereas *AtMYB37* gene (a *PsMYB91* homologous gene) can reduce the impact of salt stress by regulating membrane lipid peroxidation [64]. Transcriptome analysis revealed that the expression of *PsMYB91* and *PsMYB95* was upregulated after cold stress, indicating that these genes may be related to cold resistance in plants. Our results further support the hypothesis that the functions of homologous genes may differ between herbaceous and woody plants.

The activities and functions of genes can be deduced from their expression patterns [65]. The *MYB* gene is involved in the development and biosynthesis of plant pistils [1, 66], as well as the development of roots and leaves [67]. AtMYB106 (which interacted with PsMYB82 in the protein interaction analysis) regulates the flowering time of *A. thaliana* [68]. AtMYB16 (which interacted with PsMYB9) and AtMYB106 are associated with the formation of waxy keratin layers in reproductive organs and trichomes, which can protect cells from dehydration, block pathogen attacks, and prevent organ

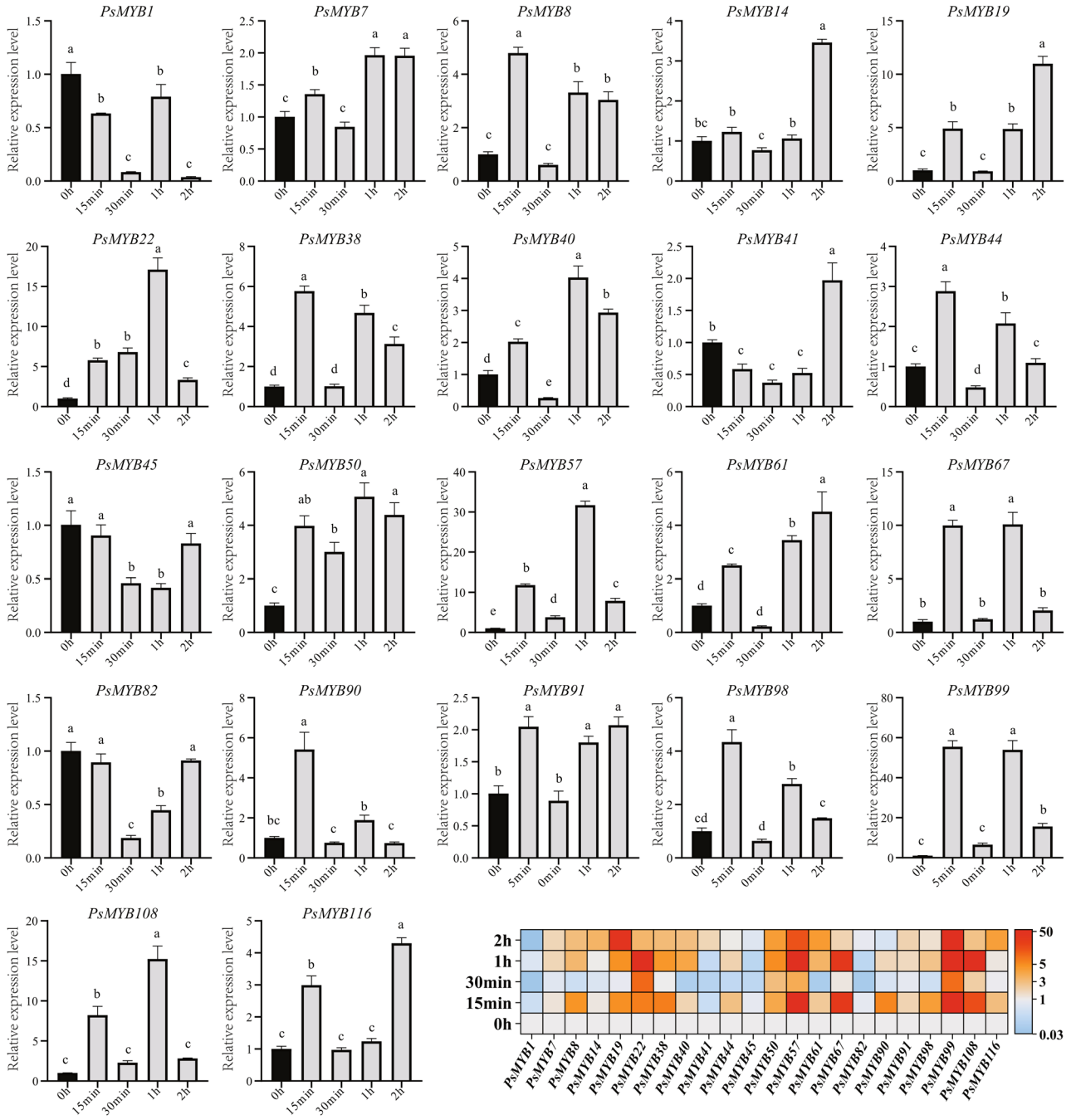

Fig. 12 Expression of 22 *PsMYBs* under different durations of cold stress. Results are the average of three replicates. The vertical bars indicate the standard deviation, and different letters over the columns denote significant differences ($p < 0.05$)

fusion during development [61]. The results of this study showed that *PsMYBs* were expressed in different parts of *P. sibirica*, indicating their essential roles in plant development. In the present study, the expression of *PsMYB9* and *PsMYB82* was significantly upregulated in the pistil, whereas that of *PsMYB41* was significantly upregulated in the leaves, confirming that these gene participate in plant development.

Frost can affect the survival of the *P. sibirica* reproductive organs in early spring, thereby seriously affecting the yield and quality of the fruit and causing significant economic losses [69]. Therefore, we conducted a study on the *MYB* gene family of *P. sibirica* and screened 22 *PsMYBs* that may be related to cold response. Gene expression analysis was conducted on the pistils of *P. sibirica* treated at low temperatures, and specific expression patterns are shown. The *RmMYB108*, *MbMYB108*, and *OsMYB2*

genes (all homologous to *PsMYB57*) exhibit a positive feedback to the cold response in *Rosa multiflora*, *Malus. baccata*, and *Oryza sativa*, respectively [70–72]. In the present study, the expression of *PsMYB57* increased significantly (by 30 folds) after 1 h of cold stress. When comparing the transcriptomes of the cold-sensitive and cold-resistant clones, *PsMYB57* expression was also significantly upregulated in the cold-resistant plant, further confirming that this gene is involved in cold response. In the homologous genes *PsMYB82* and *PsMYB38*, the latter contains *cis*-acting elements that respond to cold stress. *PsMYB38* showed an increasing expression trend in the transcriptome comparison and positive fluorescence in the quantitative analysis under low-temperature exposure, indicating its positive response to cold stress. However, the homologous gene *PsMYB82* did not show a positive response in the fluorescence results, a phenomenon that should be further studied in the future. *MbMYB4* (a *PsMYB41* homologous gene) and *AtMYB41* (a *PsMYB67* homologous gene) exhibited positive feedback to cold stress in *M. baccata* and *A. thaliana*, respectively [40, 73]. Although *PsMYB41* and *PsMYB67* are also homologous genes that can be upregulated under cold stress, the qRT-PCR results showed that *PsMYB67* was more highly expressed after cold stress treatment than *PsMYB41* was. *SaMYB98* (a *PsMYB90* homologous gene) from *Santalum album* is significantly involved in cold resistance regulation [74]. *PsMYB90* also expression under 15 min of cold stress. *PsMYB99* responded quickly to the cold stimulus, with its expression level increasing by 55 times that of the, control level after 15 min of low-temperature treatment, making it the gene that responded the fastest and showed the highest increase in expression under cold stress. PsMYB99 interacted with *AtAS1*, which regulates the development of sepals and petal primordia [75]. This study found that the expression level of *PsMYB99* was also high in the pistil, suggesting that it may be a key frost resistance gene during flowering.

Conclusions

The *R2R3-MYB* gene family plays an essential role in responding to abiotic stress in plants. In this study, we identified 116 *R2R3-MYB* gene members in *P. sibirica* and analyzed the characteristics of each gene. The functional responses of the *PsMYBs* to abiotic stress, including their responses to hormones such as jasmonate, ABA and gibberellins as well as their environmental responses to drought and low temperature, were predicted on the basis of their *cis*-acting elements, GO annotations, and protein interactions. Several candidate genes (e.g., *PsMYB99*) that may be related to cold resistance during the flowering stage of *P. sibirica* were identified through transcriptome and qRT-PCR analyses. These findings have improved our understanding of the functional characteristics of the *PsMYB* gene family and provide a theoretical basis for subsequent research on the gene molecules. Importantly, the results provide a genetic basis for cultivating high-yield and high-quality *P. sibirica* varieties with cold resistance capability.

Materials and methods

Identification and sequence analysis of *PsMYBs*

P. sibirica genome data (*Prunus sibirica F106 Genome v1.0*) were downloaded from the *Rosaceae* genome database (https://www.rosaceae.org) [76]. Overall, 126 *A. thaliana* MYB protein sequences were downloaded from The Arabidopsis Information Resource (TAIR) database (https://www.arabidopsis.org/) [77]. The Hidden Markov Model (HMM) of the conserved MYB domain (PF00249) was obtained from the Pfam database (http://pfam.xfam.org/) [72]. HMMER 3.0 software (http://hmmer.janelia.org/) was used to search for MYB domain-containing *P. sibirica* protein sequences, using the default parameters (E-value$\leq 1\times10^{-5}$). All candidate PsMYB proteins were assessed for secondary alignment using the Pfam database (https://www.ebi.ac.uk/interpro/entry/pfam/) [78] and NCBI-CDD (https://www.ncbi.nlm.nih.gov/) [79] to validate the MYB domains and further ensure the presence and completeness of the conserved *R2R3-MYB* domain. The physicochemical properties of the proteins, number of amino acids (aa), molecular weight (MW), isoelectric point (pI), instability index (II), aliphatic index (AI), Grand average of hydropathicity (GRAVY) were calculated using ExPASy (https://www.expasy.org/) [80] and subcellular localization was predicted using Plant-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant/) [81].

Multiple sequence alignment, gene structure, and motif conservation analyses

Multiple sequence alignments of the PsMYB domains were performed using Jalview 2.11.3.2 with default parameters. A multiple alignment text file was used to generate the sequence logos of *R2R3-MYB* domain repeats using WebLogo (https://weblogo.berkeley.edu/logo.cgi) [82]. TBtools v2.038 software was used to visualize the intron exon structure analysis results, and the gene structure display server 2.0 (https://gsds.gao-lab.org/) [83] was used for secondary proofreading. The online program MEME (https://meme-suite.org/meme/tools/meme) [84] was used to analyze conserved PsMYB motifs, the results of which were also visualized using TBtools software.

Phylogenetic tree analysis of PsMYBs

A phylogenetic tree was constructed in MEGA 11 using 126 AtMYB proteins of *A. thaliana* and the 116 PsMYB proteins of *P. sibirica*. Default parameter values and the

neighbor-joining (NJ) method with 1000 bootstrap iterations were used. The online beautification tool ChiPlot (https://www.chiplot.online/index.html) was used for beautification and classification of the phylogenetic tree on the basis of published *A. thaliana* classifications and *P. sibirica* motifs.

Chromosomal location and synteny analyses

Detailed information on the chromosome location of each *PsMYB* was obtained from the *P. sibirica* genome annotation analysis. The Multiple Collinearity Scan (MCScanX) toolkit of TBtools was used to illustrate the syntenic relationships of orthologous *MYB* genes between *P. sibirica* and *A. thaliana*, *P. mume*, *P. persica*, *P. salicina*, and *P. avium*. Advanced Circos of TBtools was used to determine the syntenic relationships of paralogous *MYB* genes in *P. sibirica*. The Ka-Ks calculation tool was used to calculate the Ks, Ka, and Ka/Ks ratio for the *MYB* duplicate genes in *P. sibirica*.

Analysis of *cis* -acting element and GO annotation analyses

Here, GXF Sequences Extract in TBtools was used to extract 2000 bp of the DNA sequence upstream of the *PsMYB* coding sequence,, and the gene sequences were submitted to the PlantCARE database (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [85] to predict *cis*-acting elements. Eggnog (http://eggnog5.embl.de/) [86] was used for the GO annotation analysis, and the results were visualized using WeGo (https://wego.genomics.cn/) [87].

Prediction of the protein–protein interaction network

All PsMYB proteins sequences were submitted to STRING (version 12.0; http://string-db.org), with *A. thaliana* chosen as the reference organism. After BLAST analysis, the orthologous genes of *A. thaliana* with the highest scores were used to construct a protein-protein interaction network.

Transcriptome data analysis

On the basis of the identified *PsMYBs*, transcriptome data from the pistils of the cold-tolerant "NO.453" and cold-sensitive "NO.371" clones were downloaded from the NCBI (National Center for Biotechnology Information) database (GEO number: GSE204685, https://www.ncbi.nlm.nih.gov/geo/). The gene expression patterns were detected at -4℃ for 0 h, 15 min, 30 min, 1 h, and 2 h. A Chiplot was used to construct a heatmap of visualized differential expression levels.

Plant materials and stress treatments

P. sibirica clone "Shanxing5" from the National Forest Germplasm Resource Preservation Repository for *Prunus* species at Shenyang Agricultural University (Kazuo, Liaoning, China) was used as the experimental material. Samples of healthy *P. sibirica* clone plant roots, petals, stems, pistils, and leaves were obtained from natural environments. The *P. sibirica* clones were placed in an artificial frost chamber at −4 °C, and the pistils were collected at 0 h, 15 min, 30 min, 1 h, and 2 h after low temperature exposure. The sampling and stress treatments were performed in triplicates. All samples were stored in liquid nitrogen at −80 °C.

PCR is method used to detect and quantify RNA or DNA molecules in samples. In this study, the qRT-PCR 2× Universal SYBR Green Fast qPCR Mix was used for the amplification. The PCR cycling program was as follows: 95 °C for 3 min; 40 cycles at 95 °C for 5 s, 60 °C for 30 s, and 95 °C for 15 s; and final 95 °C for 15 s. The experiment included three biological replicates. The qRT-PCR experiment in this article standardized the expression levels of *PsMYBs* using the commonly used *P. sibirica* internal reference gene *18S rRNA*, in order to analyze the differences in gene expression [69, 88], and the relative expression levels were determined using the $2^{-\Delta\Delta CT}$ method [64]. Additionally, the roots, petals, stems, pistils, and leaves of *P. sibirica* were collected to determine the expression patterns of the *PsMYB* genes in the different plant parts using qRT-PCR. The expression levels in the roots were used as the reference. Moreover, qRT-PCR to determine the gene expression levels in the pistils after 15 min, 30 min, 1 h, and 2 h of low-temperature stress at −4 °C. Pistils not subjected to cold stress were used as the control.

Abbreviations

aa	Amino acid
ABA	Abscisic acid
CAT	Catalase
DBD	DNA-binding domain
EAA	Essential amino acid
GO	Gene Ontology
HTH	Helix–turn–helix
HMM	Hidden Markov Model
pI	Isoelectric point
II	Instability index
AI	Aliphatic index
GRAVY	Grand average of hydropathicity
Ka	Non-synonymous substitution rate
Ks	Synonymous substitution rate
Ka/Ks	Ratio of the non-synonymous substitution rate to the synonymous substitution
LTR	Low-temperature response
MeJA	Methyl jasmonate
MEME	Multiple Em for Motif Elicitation
MW	Molecular weight
MYB	Myeloblastosis
POD	Peroxidase
SOD	Superoxide dismutase
TAIR	The Arabidopsis Information Resource
TF	Transcription factor

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10868-0.

Supplementary Material 1: Fig S1- Protein motif 1- motif 10.

Supplementary Material 2: Table S1- Annotations for the *cis*-acting elements.

Supplementary Material 3: Table S2- Gene Ontology (GO) annotation of *PsMYBs*.

Supplementary Material 4: Table S3- Protein–protein interactions of PsMYBs.

Supplementary Material 5: Table S4- Differential gene expression in the pistil transcriptome.

Supplementary Material 6: Table S5- Quantitative real-time PCR candidate gene information.

Supplementary Material 7: Table S6- Primer sequences for the *PsMYB* genes and internal reference gene (*18S rRNA*).

Acknowledgements
Not applicable.

Author contributions
X.Z. and Q.L. conducted experiments, analyzed data and visualized, and wrote the manuscript. Q.L. obtains funding, review and edit the manuscript. S.W., H.Z., and Y.Z. participated in the experiment and organized the data. S.D., J.C., and Q.S. are responsible for software provision and experimental verification. All authors read and approved the final manuscript.

Funding
This research was funded by the China Postdoctoral Science Foundation (2022MD723804) and Scientific Research Fund of Liaoning Provincial Education Department (LJKMZ20221056).

Data availability
Prunus sibirica F106 Genome v1.0 genome data were downloaded from the Rosaceae genome database (https://www.rosaceae.org). The AtMYB sequences were obtained from The Arabidopsis Information Resource (TAIR) database (https://www.arabidopsis.org/). The Hidden Markov Model (HMM) of the conserved MYB domain (PF00249) was obtained from the Pfam database (http://pfam.xfam.org/). The online program MEME (https://meme-suite.org/meme/tools/meme) was used to analyze conserved PsMYB motifs. The PlantCARE database (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to predict cis-acting elements. The STRING database (http://string-db.org) to prediction of the protein–protein interaction network.All PsMYBs sequences were submitted to. Transcriptome data was downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/geo/) (NCBI GEO number: GSE204685). The datasets analysed during this study are included in this published article and its supplementary information files.

Declaration

Ethics approval and consent to participate
Plant material used in the study complies with relevant institutional, national, and international guidelines and legislation. Plant material used is not wild species and also permission for the plant material is not required.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 6 June 2024 / Accepted: 4 October 2024
Published online: 14 October 2024

References

1. Wang Y, Zhou H, He Y, Shen X, Lin S, Huang L. MYB transcription factors and their roles in the male reproductive development of flowering plants. Plant Sci. 2023;335:111811.
2. Li J, Han G, Sun C, Sui N. Research advances of MYB transcription factors in plant stress resistance and breeding. Plant Signal Behav. 2019;14(8):1613131.
3. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in *Arabidopsis*. Trends Plant Sci. 2010;15(10):573–81.
4. Hou XJ, Li SB, Liu SR, Hu CG, Zhang JZ. Genome-wide classification and evolutionary and expression analyses of citrus MYB transcription factor families in sweet orange. PLoS ONE. 2014;9(11):e112375.
5. Ogata K, Kanei Ishii C, Sasaki M, Hatanaka H, Nagadoi A, Enari M, et al. The cavity in the hydrophobic core of myb DNA-binding domain is reserved for DNA recognition and *trans*-activation. Nat Struct Biol. 1996;3(2):178–87.
6. Klempnauer KH, Gonda TJ, Bishop JM. Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. Cell. 1982;31(2):453–63.
7. Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H. The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. 1987;6(12):3553–8.
8. Liu H, Xiong JS, Jiang YT, Wang L, Cheng ZMM. Evolution of the *R2R3-MYB* gene family in six *Rosaceae* species and expression in woodland strawberry. J Integr Agric. 2019;18(12):2753–70.
9. Chen G, He W, Guo A, Pan J. Genome-wide identification, classification and expression analysis of the MYB transcription factor family in petunia. Int J Mol Sci. 2021;22(9):4838.
10. Long F, Wu H, Li H, Zuo W, Ao Q. Genome-wide analysis of MYB transcription factors and screening of MYBs involved in the red color formation in *Rhododendron Delavayi*. Int J Mol Sci. 2023;24(5):4641.
11. Zhang Z, Liu Z, Wu H, Xu Z, Zhang H, Qian W, et al. Genome-wide identification and characterization of *MYB* gene family and analysis of its sex-biased expression pattern in *Spinacia oleracea* L. Int J Mol Sci. 2024;25(2):795.
12. Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in Arabidopsis thaliana. Curr Opin Plant Biol. 2001;4(5):447–56.
13. Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM. Expansion and diversification of the *Populus R2R3-MYB* family of transcription factors. Plant Physiol. 2009;149(2):981–93.
14. Jia L, Clegg MT, Jiang T. Evolutionary dynamics of the DNA-binding domains in putative *R2R3-MYB* genes identified from rice subspecies *indica* and *japonica* genomes. Plant Physiology. 2004;134(2):575–585.
15. Xu X, Yang Y, Zhang T, Zhang Y, Tong H, Yuan H. Systematic analysis of the R2R3-MYB transcription factor gene family in *Stevia rebaudiana*. Ind Crops Prod. 2024;210:118123.
16. Zhou F, Chen Y, Wu H, Yin T. Genome-wide comparative analysis of *R2R3 MYB* gene family in *Populus* and *Salix* and identification of male flower bud development-related genes. Front Plant Sci. 2021;12:721558.
17. Li J, Guo S, Min Htwe Y, Sun X, Zhou L, Wang F, et al. Genome-wide identification, classification and expression analysis of *MYB* gene family in coconut (*Cocos nucifera* L). Front Plant Sci. 2024;14:1263595.
18. Liu X, Wu Z, Feng J, Yuan G, He L, Zhang D, et al. A novel *R2R3-MYB* gene *LoMYB33* from lily is specifically expressed in anthers and plays a role in pollen development. Front Plant Sci. 2021;12:730007.
19. Zepeda B, Marcelis LFM, Kaiser E, Verdonk JC. Petunia as a model for MYB transcription factor action under salt stress. Front Plant Sci. 2023;14:1286547.
20. Liu M, Yang H, Fan S, Guo B, Dai L, Wang L, et al. Genome-wide identification and expression analysis of the *R2R3-MYB* gene family in rubber trees. Forests. 2023;14(4):710.
21. Ma X, Yu YN, Jia JH, Li QH, Gong ZH. The pepper MYB transcription factor *CaMYB306* accelerates fruit coloration and negatively regulates cold resistance. Sci Hort. 2022;295:110892.
22. Li W, Zhong J, Zhang L, Wang Y, Song P, Liu W, et al. Overexpression of a *Fragaria vesca* MYB transcription factor gene (*FvMYB82*) increases salt and cold tolerance in *Arabidopsis thaliana*. Int J Mol Sci. 2022;23(18):10538.
23. So K, Wang J, Sun S, Che H, Zhang Y. Comprehensive analysis of *MYB* gene family and their expression under various stress conditions in *Lilium Pumilum*. Sci Hort. 2024;327:112764.
24. Li Y, Zhang J, Wang S, Liu Y, Yang M, Huang Y. Genome-wide identification of the *Pyrus R2R3-MYB* gene family and *PhMYB62* regulation analysis in *Pyrus hopeiensis* flowers at low temperature. Int J Biol Macromol. 2024;257:128611.

25. Yin M, Wuyun T, Jiang Z, Zeng J. Amino acid profiles and protein quality of siberian apricot (*Prunus Sibirica* L.) kernels from Inner Mongolia. J Forestry Res. 2020;31(4):1391–7.
26. Liu R, Chen J, Zhang Y, Wang P, Kang Y, Li B, et al. Physiological and biochemical characteristics of *Prunus Sibirica* during flowering. Sci Hort. 2023;321:112358.
27. Shu L, Li LH, Jiang YQ, Yan JL. Advances in membrane-tethered NAC transcription factors in plants. Plant Sci. 2024;342:112034.
28. Baillo EH, Kimotho RN, Zhang Z, Xu. Transcription factors associated with abiotic and biotic stress tolerance and their potential for crops improvement. Genes. 2019;10(10):771.
29. Liu H, Tang X, Zhang N, Li S, Si H. Role of bZIP transcription factors in plant salt stress. Int J Mol Sci. 2023;24(9):7893.
30. Sayou C, Nanao MH, Jamin M, Pose D, Thevenon E, Gregoire L, et al. A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. Nat Commun. 2016;7:11222.
31. Cao Y, Jia H, Xing M, Jin R, Grierson D, Gao Z, et al. Genome-wide analysis of *MYB* gene family in Chinese bayberry (*Morella rubra*) and identification of members regulating flavonoid bosynthesis. Front Plant Sci. 2021;12:691384.
32. Li C, Ng CKY, Fan LM. MYB transcription factors, active players in abiotic stress signaling. Environ Exp Bot. 2015;114:80–91.
33. Liu C, Hao J, Qiu M, Pan J, He Y. Genome-wide identification and expression analysis of the MYB transcription factor in Japanese plum (*Prunus salicina*). Genomics. 2020;112(6):4875–86.
34. Reyes JL, Campos F, Wei H, Arora R, Yang Y, Karlson DT, et al. Functional dissection of hydrophilins during in vitro freeze protection. Plant Cell Environ. 2008;31(12):1781–90.
35. Saikumar P, Murali R, Reddy EP. Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions. Proceedings of the National Academy of Sciences. 1990;87(21):8452–8456.
36. Yang J, Zhang B, Gu G, Yuan J, Shen S, Jin L, et al. Genome-wide identification and expression analysis of the *R2R3-MYB* gene family in tobacco (*Nicotiana tabacum* L). BMC Genomics. 2022;23(1):432.
37. Yuan Y, Wang Q, Tong B, Wang Z, Liu D, Yan Q, et al. Identification of R2R3-MYB members in Lonicera japonica: insights into the positive role of Lj5. MYB61 and Lj4.MYB16 in glandular trichome density. Ind Crops Prod. 2024;216:118738.
38. Abdulrahman SS, Daştan SD, Shahbaz SE, Selamoglu Z. Phylogenetic analysis of *Prunus* genus using nuclear and chloroplast gene markers as a bioorganic structure profiling. J Mol Struct. 2023;1284:135300.
39. Zhang P, Wang R, Yang X, Ju Q, Li W, Lü S, et al. The *R2R3-MYB* transcription factor *AtMYB49* modulates salt tolerance in *Arabidopsis* by modulating the cuticle formation and antioxidant defence. Plant Cell Environ. 2020;43(8):1925–43.
40. Cominelli E, Sala T, Calvi D, Gusmaroli G, Tonelli C. Over-expression of the *Arabidopsis AtMYB41* gene alters cell expansion and leaf surface permeability. Plant J. 2008;53(1):53–64.
41. Lippold F, Sanchez DH, Musialak M, Schlereth A, Scheible W-R, Hincha DK, et al. *AtMyb41* regulates transcriptional and metabolic responses to osmotic stress in *Arabidopsis*. Plant Physiol. 2009;149(4):1761–72.
42. Wang A, Liang K, Yang S, Cao Y, Wang L, Zhang M, et al. Genome-wide analysis of MYB transcription factors of *Vaccinium corymbosum* and their positive responses to drought stress. BMC Genomics. 2021;22(1):565.
43. Lee SB, Kim HU, Suh MC. *MYB94* and *MYB96* additively activate cuticular wax biosynthesis in *Arabidopsis*. Plant and Cell Physiology. 2016;57(11):2300–2311.
44. Shi L, Chen Y, Hong J, Shen G, Schreiber L, Cohen H, et al. *AtMYB31* is a wax regulator associated with reproductive development in *Arabidopsis*. Planta. 2022;256(2):1–15.
45. Castorina G, Domergue F, Chiara M, Zilio M, Persico M, Ricciardi V, et al. Drought-responsive *ZmFDL1/MYB94* regulates cuticle biosynthesis and cuticle-dependent leaf permeability. Plant Physiol. 2020;184(1):266–82.
46. Agarwal M, Hao Y, Kapoor A, Dong CH, Fujii H, Zheng X, et al. A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. J Biol Chem. 2006;281(49):37636–45.
47. Wang X, Ding Y, Li Z, Shi Y, Wang J, Hua J, et al. *PUB25* and *PUB26* promote plant freezing rolerance by degrading the cold signaling negative regulator *MYB15*. Dev Cell. 2019;51(2):222–35.
48. Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, et al. Differential regulation of closely related *R2R3-MYB* transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. Plant J. 2007;50(4):660–77.
49. Wang YW, Hess J, Slot JC, Pringle A. De novo gene birth, horizontal gene transfer, and gene duplication as sources of new gene families associated with the origin of symbiosis in *Amanita*. Genome Biol Evol. 2020;12(11):2168–82.
50. Ma R, Huang W, Hu Q, Tian G, An J, Fang T, et al. Tandemly duplicated *MYB* genes are functionally diverged in the regulation of anthocyanin biosynthesis in soybean. Plant Physiol. 2024;194(4):2549–63.
51. Du J, Zhang Q, Hou S, Chen J, Meng J, Wang C, et al. Genome-wide identification and analysis of the *R2R3-MYB* gene family in *Theobroma cacao*. Genes. 2022;13(9):1572.
52. Shi S, Li D, Li S, Zhao N, Liao J, Ge H, et al. Genome-wide analysis of *R2R3-MYB* genes and functional characterization of *SmMYB75* in eggplant fruit implications for crop improvement and nutritional enhancement. Int J Mol Sci. 2024;25(2):1163.
53. Harouaka D, Engle RE, Wollenberg K, Diaz G, Tice AB, Zamboni F et al. Diminished viral replication and compartmentalization of hepatitis C virus in hepatocellular carcinoma tissue. Proceedings of the National Academy of Sciences. 2016;113(5):1375–1380.
54. He Y, Liu W, Wang J. Assembly and comparative analysis of the complete mitochondrial genome of *Trigonella foenum-graecum* L. BMC Genomics. 2023;24(1):1–19.
55. Zou X, Sun H. DOF transcription factors: specific regulators of plant biological processes. Front Plant Sci. 2023;14:1044918.
56. Zelicourt AD, Colcombet J, Hirt H. The role of MAPK modules and ABA during abiotic stress signaling. Trends Plant Sci. 2016;21(8):677–85.
57. Gul N, Masoodi KZ, Ramazan S, Mir JI, Aslam S. Study on the impact of exogenously applied methyl jasmonate concentrations on *Solanum lycopersicum* under low temperature stress. BMC Plant Biol. 2023;23(1):1–12.
58. Han Z, Zhang C, Zhang H, Duan Y, Zou Z, Zhou L, et al. CsMYB transcription factors participate in jasmonic acid signal transduction in response to cold stress in tea plant (*Camellia sinensis*). Plants. 2022;11(21):2869.
59. Wang T, Tohge T, Ivakov A, Mueller-Roeber B, Fernie AR, Mutwil M, et al. Salt-related *MYB1* coordinates abscisic acid biosynthesis and signaling during salt stress in *Arabidopsis*. Plant Physiol. 2015;169(2):1027–41.
60. Zhang Y, Xu J, Li R, Ge Y, Li Y, Li R. Plants' response to abiotic stress: mechanisms and strategies. Int J Mol Sci. 2023;24(13):10915.
61. Oshima Y, Shikata M, Koyama T, Ohtsubo N, Mitsuda N, Ohme-Takagi M. MIXTA-like transcription factors and WAX INDUCER1/SHINE1 coordinately regulate cuticle development in *Arabidopsis* and *Torenia fournieri*. Plant Cell. 2013;25(5):1609–24.
62. Zhou L, Yarra R, Yang Y, Liu Y, Yang M, Cao H. The oil palm *R2R3-MYB* subfamily genes *EgMYB111* and *EgMYB157* improve multiple abiotic stress tolerance in transgenic *Arabidopsis* plants. Plant Cell Rep. 2022;41(2):377–93.
63. Uemura Y, Kimura S, Ohta T, Suzuki T, Mase K, Kato H, et al. A very long chain fatty acid responsive transcription factor, *MYB93*, regulates lateral root development in *Arabidopsis*. Plant J. 2023;115(5):1408–27.
64. Li Y, Tian B, Wang Y, Wang J, Zhang H, Wang L, et al. The transcription factor *MYB37* positively regulates photosynthetic inhibition and oxidative damage in *Arabidopsis* leaves under salt stress. Front Plant Sci. 2022;13:943153.
65. Sabir IA, Manzoor MA, Shah IH, Liu X, Zahid MS, Jiu S, et al. MYB transcription factor family in sweet cherry (*Prunus avium* L.): genome-wide investigation, evolution, structure, characterization and expression patterns. BMC Plant Biol. 2022;22(1):1–20.
66. Iqbal S, Pan Z, Hayat F, Bai Y, Coulibaly D, Ali S, et al. Comprehensive transcriptome profiling to identify genes involved in pistil abortion of Japanese apricot. Physiol Mol Biology Plants. 2021;27(6):1191–204.
67. Liu Y, Ma D, Constabel CP. CRISPR/Cas9 disruption of *MYB134* and *MYB115* in transgenic poplar leads to differential reduction of proanthocyanidin synthesis in roots and leaves. Plant Cell Physiol. 2023;64(10):1189–203.
68. Hong L, Niu F, Lin Y, Wang S, Chen L, Jiang L. *MYB106* is a negative regulator and a substrate for $CRL3^{BPM}$ E3 ligase in regulating flowering time in *Arabidopsis thaliana*. J Integr Plant Biol. 2021;63(6):1104–19.
69. Liu Q, Wang S, Wen J, Chen J, Sun Y, Dong S. Genome-wide identification and analysis of the WRKY gene family and low-temperature stress response in *Prunus Sibirica*. BMC Genomics. 2023;24(1):358.
70. Dong J, Cao L, Zhang X, Zhang W, Yang T, Zhang J, et al. An R2R3-MYB transcription factor *RmMYB108* responds to chilling stress of *Rosa multiflora* and conferred cold tolerance of *Arabidopsis*. Front Plant Sci. 2021;12:696919.
71. Yang A, Dai X, Zhang WH. A R2R3-type *MYB* gene, *OsMYB2*, is involved in salt, cold, and dehydration tolerance in rice. J Exp Bot. 2012;63(7):2541–56.
72. Yao C, Li W, Liang X, Ren C, Liu W, Yang G, et al. Molecular cloning and characterization of *MbMYB108*, a *Malus baccata* MYB transcription factor gene, with

functions in tolerance to cold and drought stress in transgenic *Arabidopsis thaliana*. Int J Mol Sci. 2022;23(9):4846.
73. Yao C, Li X, Li Y, Yang G, Liu W, Shao B, et al. Overexpression of a *Malus baccata* MYB transcription factor gene *MbMYB4* increases cold and drought tolerance in *Arabidopsis thaliana*. Int J Mol Sci. 2022;23(3):1794.
74. Tang M, Liu L, Hu X, Zheng H, Wang Z, Liu Y, et al. Genome-wide characterization of *R2R3-MYB* gene family in *Santalum album* and their expression analysis under cold stress. Front Plant Sci. 2023;14:1142562.
75. Xu B, Li Z, Zhu Y, Wang H, Ma H, Dong A et al. *Arabidopsis* genes *AS1*, *AS2*, and *JAG* negatively regulate boundary-specifying genes to promote sepal and petal development. Plant Physiology. 2008;146(2):566.
76. Jung S, Lee T, Cheng C-H, Buble K, Zheng P, Yu J, et al. 15 years of GDR: New data and functionality in the genome database for *Rosaceae*. Nucleic Acids Res. 2019;47(D1):D1137–45.
77. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al. The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 2012;40(D1):D1202–10.
78. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: the protein families database in 2021. Nucleic Acids Res. 2021;49(D1):D412–9.
79. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 2020;48(D1):D265–8.
80. Duvaud S, Gabella C, Lisacek F, Stockinger H, Ioannidis V, Durinx C. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. Nucleic Acids Res. 2021;49(W1):W216–27.
81. Chou KC, Shen HB. Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. PLoS ONE. 2010;5(6):e11335.
82. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. Genome Res. 2004;14(6):1188–90.
83. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31(8):1296–7.
84. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. Nucleic Acids Res. 2015;43(W1):W39–49.
85. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van De Peer Y, et al. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.
86. Huerta Cepas J, Szklarczyk D, Heller D, Hernández Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019;47(D1):D309–14.
87. Ye J, Zhang Y, Cui H, Liu J, Wu Y, Cheng Y, et al. WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic Acids Res. 2018;46(W1):W71–5.
88. Liu Q, Wen J, Wang S, Chen J, Sun Y, Liu Q, et al. Genome-wide identification, expression analysis, and potential roles under low-temperature stress of bHLH gene family in *Prunus Sibirica*. Front Plant Sci. 2023;14:1267107.

Publisher's note