Molecular Plant Breeding 2022, Vol.13, No.23, 1-10
http://genbreedpublisher.com/index.php/mpb

Research Article Open Access

Analysis of Transcriptome Sequencing and MYB Transcription Factor Family in *Rhododendron lapponicum*

Xinping Jia, Xiaoqing Liu, Yanming Deng, Chang Li, Jiale Su
Jiangsu Academy of Agricultural Sciences, Institute of Leisure Agriculture, Nanjing, 210014, China
Co-corresponding authors email: pingdaya@163.com
Molecular Plant Breeding, 2022, Vol.13, No.23 doi: 10.5376/mpb.2022.13.0023
Received: 09 Aug., 2022
Accepted: 15 Aug., 2022
Published: 23 Aug., 2022

Preferred citation for this article:
Jia X.P., Liu X.Q., Deng Y.M., Li C., and Su J.L., 2022, Analysis of transcriptome sequencing and MYB transcription factor family in *Rhododendron lapponicum*, Molecular Plant Breeding, 13(23): 1-10 (doi: 10.5376/mpb.2022.13.0023)

Abstract MYB transcription factor is the largest family of transcription factors in plants, and they are involved in the regulation of plant growth and development, secondary metabolism, adversity stress and other biological processes. So far, there is no the study on the MYB transcription factor of *Rhododendron lapponicum*. In this study, the *Rhododendron lapponicum* variety 'Fuli Jinling' transcriptome was generated by SMRT sequencing technology. A total of 15.37 Gb data was obtained, and 75 002 transcript sequences were obtained by removing redundancy. 71 155, 33 653 and 30 359 transcripts were assigned to the Nr, GO and COG databases, respectively. Based on the transcriptome sequencing data, 64 transcription factor gene families were identified, including 220 MYB genes. According to the structural characteristics, the MYB gene is divided into four categories, including 1R-MYB, R2R3-MYB, R1R2R3-MYB and 4R-MYB. The amino acid sequence of MYB transcription factor contains 20 conserved elements. Phylogenetic analysis showed that the MYB genes of *Rhododendron lapponicum* could be divided into 28 subclasses. In this study, we first used SMRT sequencing technology to generate the *Rhododendron lapponicum* transcriptome. In this study, single molecule real time (SMRT) technology was used to sequence the transcriptome of the *Rhododendron lapponicum* variety 'Fuli Jinling', and the obtained transcript sequences were functionally annotated and classified, and 220 *MYB* genes for bioinformatics analysis, these related results have certain reference significance.
Keywords *Rhododendron lapponicum*; Transcriptome; MYB transcription factor

Rhododendron simsii Planch., also known as Yingshanhong, Shanzhizhu, Shanshiliu etc. in Chinese, is an evergreen shrub of Rhododendron in Ericaceae. It is a worlds' famous ornamental flower and one of the ten traditional famous flowers in China. There are 960 species of Rhododendron in the world, and China is the country with the most abundant species of Rhododendron, with over 570 species of 6 subgenera, and is the main distribution center of *Rhododendron lapponicum* resources in the world (Huang et al., 2005). *Rhododendron lapponicum* belongs to the Family Rhododenaceae, which is a *Subgen.* Hymenanthes plant with brilliant colors and bright evergreen leaves. It is suitable for potted plants and landscaping with high ornamental value and economic value (Liu et al., 2011). At present, researches on *Rhododendron lapponicum* are mainly focused on morphological classification, adversity stress, cultivation, breeding and reproduction techniques (Li et al., 2018; Chen et al., 2019).

Transcriptome sequencing can quickly obtain the overall situation of gene expression in a specific cell or tissue of a species in a certain state, which can be used to study gene structure and function, alternative splicing and prediction of new transcripts. For species without reference genome information, transcriptome sequencing provides a new research idea for mining functional genes, alternative splicing events, and molecular marker development (Xiao et al., 2013; Li et al., 2013; Zhang et al., 2015). Transcription factor is a kind of regulatory protein with special structure, which combines with cis-acting elements of target genes to regulate RNA transcription and expression (Hobert, 2008). MYB transcription factors are the largest transcription factor family in plants and are involved in the regulation of anthocyanin synthesis, hormone signal transduction, stress response, organ development and other biological processes (Dai et al., 2007; Kranzr et al., 2010). At present, there is no

report on the MYB transcription factor of *Rhododendron lapponicum*. In this study, single molecule real time (SMRT) sequencing was used to sequence the transcriptome of *Rhododendron lapponicum* 'Fuli Jinling'. The obtained transcriptome was functionally annotated and classified. MYB transcription factor family was analyzed by bioinformatics methods. It will provide an important basis for the follow-up study on the function of *MYB* gene in *Rhododendron simsii*.

1 Results and Analysis

1.1 Statistical analysis of transcriptome data of *Rhododendron lapponicum*

The transcriptome library of *Rhododendron lapponicum* was sequenced by SMRT sequencing technology, and 15.37 Gb of sequencing data were obtained. 658 338 Reads of Insert (ROI) sequences were extracted from the original sequence with an average length of 2 216 bp. After removing the cDNA primers and polyA sequences in ROI sequence, 346 270 full-length non-chimeric sequences were obtained. SMRT Analysis software was used to cluster FLNC sequences and 105 015 transcripts were obtained. Finally, 75 002 non-redundant full-length transcripts were obtained from the high-quality transcripts and the corrected transcripts. The total length of the transcript was 180.45 Mb, with an average length of 2 406 bp. According to the sequence length distribution statistics of *Rhododendron lapponicum* transcripts, 971 transcripts <500 bp, accounting for 1.29% of the total. There were 27 481 transcripts of 500~2 000 bp, accounting for 36.64% of the total. There were 46 550 transcripts>2 000 bp, accounting for 62.07% of the total.

1.2 Functional annotation and classification of *Rhododendron lapponicum* transcripts

BLAST software was used to compare the full-length transcripts of *Rhododendron lapponicum* with the NR protein database, and 71 155 transcripts could be found in the NR database. Among the matched related species, the grape (*Vitis vinifera*) (17.71%) had the highest proportion, followed by oak (*Quercus suber*) (5.59%), walnut (*Juglans regia*) (3.47%), coffee (*Coffea Linn*) (3.42%), and other species. The transcripts were functionally classified based on GO (Gene Ontology) database, and the functional distribution characteristics of genes expressed in *Rhododendron lapponicum*s were analyzed. The results showed that 33 653 transcripts of *Rhododendron lapponicum* had GO entries, which were divided into 51 subclasses (Table 1). In the cell component classification, the number of cell parts (13 440) and cells (13 394) was the largest. In the classification of biological processes, metabolic processes (23 749) accounted for the highest proportion. In molecular functional classification, catalytic activities (18 900) and protein binding (16 707) were more numerous, while metallochaperone activity (9) and protein tags (7) were less numerous.

Comparing the transcripts of *Rhododendron lapponicum* with the Clusters of Orthologous Groups (COG) database, we found 30359 transcripts with functional information in the COG database, which were divided into 24 functional categories (Figure 1). Among them, the number of general function genes was the most (8 967), followed by transcription genes (4 862), replication, recombination and repair genes (4 722) and signal transduction mechanism genes (4 428), while the number of nuclear structure genes (27) and cell movement genes (7) were relatively small.

1.3 Identification of transcription factor families in *Rhododendron lapponicum*

Based on transcriptome sequencing data of *Rhododendron lapponicum*, 3 287 transcription factor genes were predicted, including 64 transcription factor families. C3H transcription factor family had the largest number of genes (231), followed by MYB (220), FAR1 (187), bHLH (182), C2H2 (150), GRAS (136), bZIP (130), WRKY (123), etc. (Table 2).

1.4 Analysis of MYB transcription factor family in *Rhododendron lapponicum*

According to the DNA binding domain characteristics contained in MYB transcription factors, the MYB genes of *Rhododendron lapponicum* were divided into four classes, they are, 1R-MYB, R2R3-MYB, R1R2R3-MYB and 4R-MYB, named as *SaMYB1~SaMYB220*. 157 1R-MYB transcription factors, 60 R2R3-MYB transcription factors, 1 R1R2R3-MYB transcription factors, and 2 4R-MYB transcription factors (Table 3).

Table 1 GO functional categories of *Rhododendron lapponicum* transcripts

GO annotation	GO function classification	Number
Cellular Component	Extracellular region	861
	Cell	13 394
	Nucleoid	10
	Membrane	7 904
	Virion	6
	Cell junction	634
	Extracellular matrix	8
	Membrane-enclosed lumen	353
	Macromolecular complex	3 616
	Organelle	9 656
	Extracellular matrix part	8
	extracellular region part	24
	Organelle part	4 895
	Virion part	6
	Membrane part	3 966
	Cell part	13 440
Molecular Function	Protein binding transcription factor activity	50
	Nucleic acid binding transcription factor activity	501
	Catalytic activity	18 900
	Receptor activity	194
	Guanyl-nucleotide exchange factor activity	36
	Structural molecule activity	959
	Transporter activity	2 405
	Binding	16 707
	Electron carrier activity	614
	Antioxidant activity	220
	Metallochaperone activity	8
	Enzyme regulator activity	323
	Protein tag	10
	Nutrient reservoir activity	25
	Molecular transducer activity	469
Biological Process	Reproduction	382
	Cell killing	96
	Immune system process	438
	Metabolic process	23 749
	Cellular process	19 892
	Reproductive process	1 369
	Biological adhesion	160
	Signaling	1 593
	Multicellular organismal process	2084
	Developmental process	2827
	Growth	553
	Locomotion	33
	Single-organism process	16 672
	Biological phase	136
	Rhythmic process	100
	Response to stimulus	6 515
	Localization	5 341
	Multi-organism process	1 054
	Biological regulation	5 966
	Cellular component organization or biogenesis	3 722

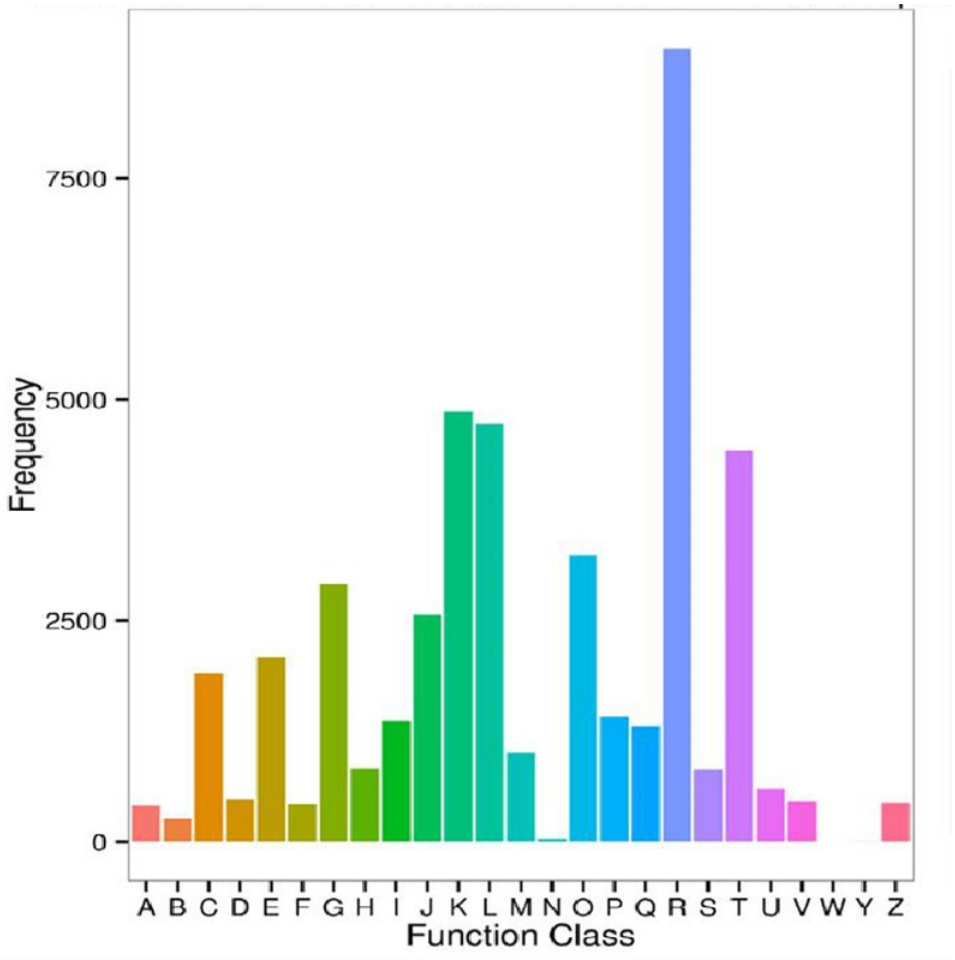

Figure 1 COG function classification of *Rhododendron lapponicum* transcripts

Note: A: RNA processing and modification; B: Chromatin structure and dynamics; C: Energy production and conversion; D: Cell cycle control, cell division, chromosome partitioning; E: Amino acid transport and metabolism; F: Nucleotide transport and metabolism; G: Carbohydrate transport and metabolism; H: Coenzyme transport and metabolism; I: Lipid transport and metabolism; J: Translation, ribosomal structure and biogenesis; K: Transcription; L: Replication, recombination and repair; M: Cell wall/membrane/envelope biogenesis; N: Cell motility; O: Posttranslational modification, protein turnover, chaperones; P: Inorganic ion transport and metabolism; Q: Secondary metabolites biosynthesis, transport and catabolism; R: General function prediction only; S: Function unknown; T: Signal transduction mechanisms; U: Intracellular trafficking, secretion, and vesicular transport; V: Defense mechanisms; W: Extracellular structure; Y: Nuclear structure; Z: Cytoskeleton

1.5 Analysis of conserved elements of MYB transcription factor in *Rhododendron lapponicum*

The amino acid sequences of the MYB transcription factor of *Rhododendron lapponicum* contained 20 conserved elements (Figure 2). Different *Rhododendron lapponicum* MYB transcription factor genes contained different numbers of conservative elements. *SaMYB156* gene contained the least number of conservative elements (1), while *SaMYB39*, *SaMYB40*, *SaMYB42*, *SaMYB43*, *SaMYB44* and *SaMYB45* genes contained the most number of conservative elements (12). The letters in the figure highly expressed the degree of conserved domain of amino acid residues at this site, and the smaller the letters were, the smaller the probability of occurrence and conserved degree of amino acids expressed at the corresponding site was.

1.6 Phylogenetic tree analysis of MYB transcription factor family in *Rhododendron lapponicum*

The amino acid sequences of 196 Arabidopsis *MYB* genes were compared with 220 *Rhododendron lapponicum MYB* genes by multi-sequence comparison, and phylogenetic tree was constructed by MEGA software. The results showed that the MYB genes of *Rhododendron lapponicum* and *Arabidopsis thaliana* were divided into 28 subgroups (Figure 3).

Table 2 Transcription factor family classification of *Rhododendron lapponicum* transcripts

No.	Pathway	Number	No.	Pathway	Number
1	C3H	231	33	BES1	30
2	MYB	220	34	NF-YC	28
3	FAR1	187	35	EIL	24
4	bHLH	182	36	NF-YA	22
5	C2H2	150	37	HB-KNOX	21
6	GRAS	136	38	BBR-BPC	21
7	MYB-related	135	39	Alfin-like	19
8	bZIP	130	40	C2C2-LSD	18
9	WRKY	123	41	zf-HD	15
10	RWP-RK	120	42	LOB	14
11	NAC	115	43	AP2/ERF-AP2	14
12	AP2/ERF-ERF	115	44	PLATZ	13
13	B3-ARF	103	45	NF-YB	13
14	B3	96	46	HB-PHD	12
15	HB-HD-ZIP	76	47	DBB	12
16	GARP-G2-like	71	48	VOZ	10
17	Trihelix	68	49	CPP	9
18	SBP	60	50	C2C2-YABBY	9
19	HB-other	58	51	Whirly	7
20	C2C2-GATA	56	52	STAT	7
21	HB-BELL	53	53	HRT	7
22	MADS-MIKC	50	54	DBP	6
23	Tify	45	55	LIM	5
24	C2C2-Dof	45	56	S1Fa-like	4
25	HSF	44	57	OFP	4
26	GeBP	44	58	NF-X1	4
27	TUB	41	59	CSD	4
28	TCP	37	60	SRS	2
29	C2C2-CO-like	36	61	GRF	2
30	GARP-ARR-B	34	62	E2F-DP	2
31	CAMTA	34	63	BSD	2
32	MADS-M-type	31	64	HB-WOX	1

Table 3 Protein sequence analysis of MYB transcription factor family

Type	Number	Protein length	Relative molecular weight
1R-MYB	157	53-1656 aa	6.13-179.54 kD
R2R3-MYB	60	137~1035 aa	15.89~114.95 kD
R1R2R3-MYB	1	577 aa	64.19 kD
4R-MYB	2	597~650 aa	67.72~73.72 kD
Total	220	53~1656 aa	6.13~179.54 kD

2 Discussion

With the rapid development of molecular biology, a new generation of high-throughput sequencing technology has been widely used in plant transcriptome research. In recent years, high-throughput sequencing technologies have been constantly innovated, such as the third-generation sequencing characterized by single-molecule sequencing-real-time single-molecule sequencing of Pacific Bioscience, Nanopore single-molecule sequencing of Oxford Nanopore Technologies, which is considered to be an ideal sequencing platform for whole-genome assembly and full-length transcript sequencing (Branton et al., 2008). The third-generation sequencing technology does not require PCR amplification, and has the characteristics of fast speed, long read length, and no PCR

amplification bias and GC bias. In the absence of reference genomes, third-generation sequencing technology has become an important method for plant gene mining, epigenetic studies, and genetic diversity analysis (Liu et al., 2017; Ardui et al., 2018; Jia et al., 2020). In this study, the transcriptome of *Rhododendron lapponicum* alpine was analyzed by high-throughput SMRT sequencing technology, and 15.37 Gb of sequencing data were obtained, and 75 002 non-redundant full-length transcripts were obtained. The average length of *Rhododendron lapponicum* transcripts obtained by SMRT sequencing (2 406 bp) was significantly higher than that of sweet potato (581 bp) (Wang et al., 2010) and sesame (629 bp) (Wei et al., 2011) and *Asplenium nidus* (936 bp) (Jia et al., 2016), indicating that SMRT sequencing technology can directly obtain complete full-length transcripts without splice, overcomes the disadvantages of short read length of second-generation sequencing technology, and provides basic data for further research on gene function of *Rhododendron lapponicum*.

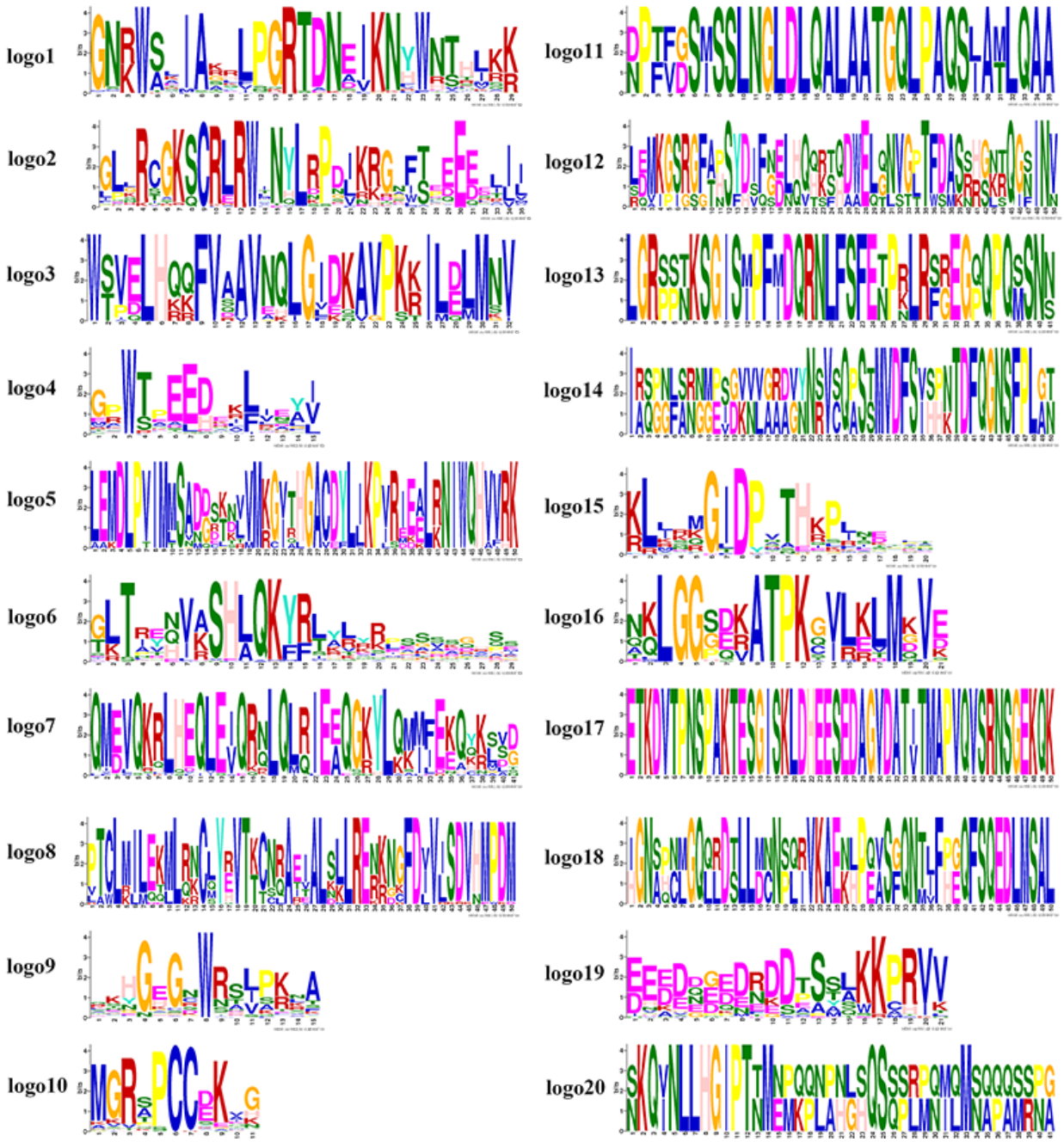

Figure 2 Conserved elements of mango MYB family proteins

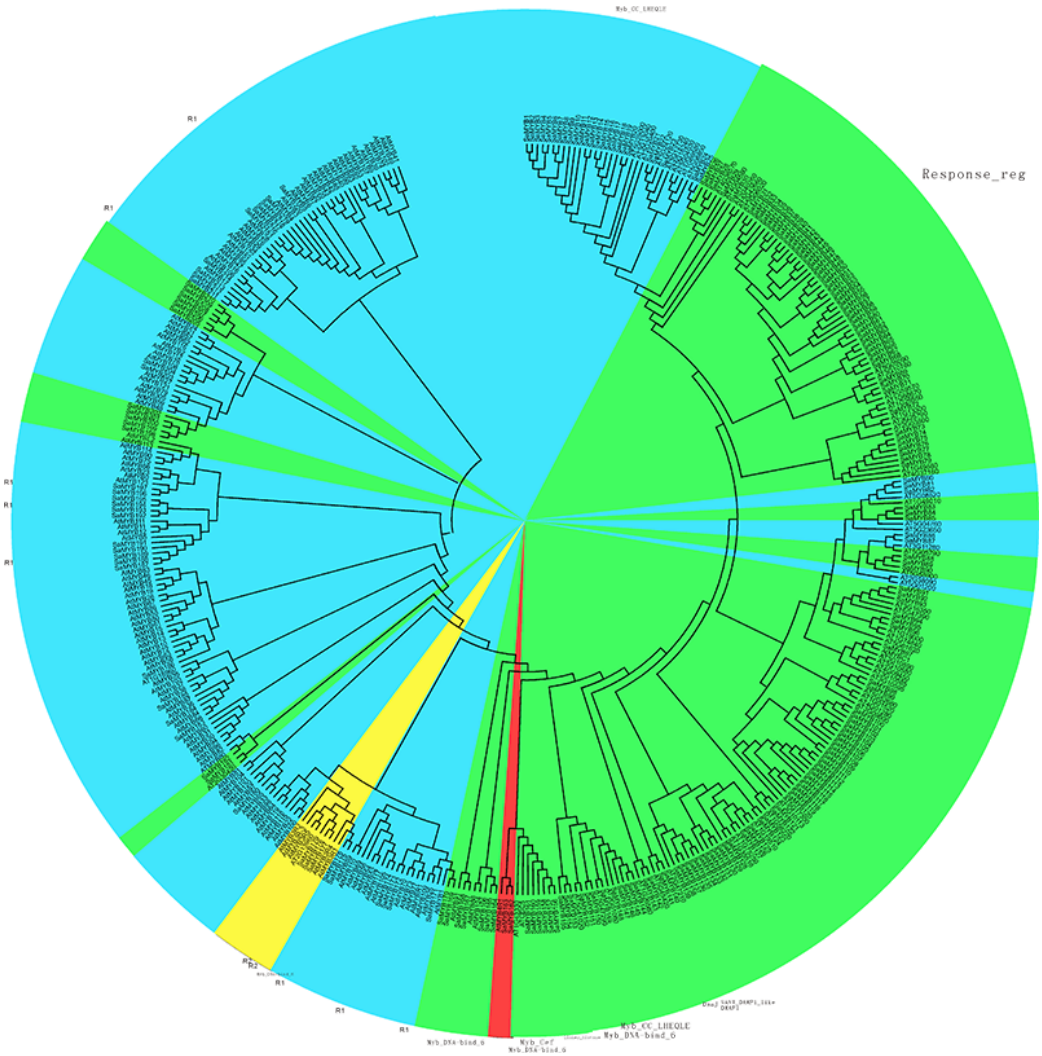

Figure 3 Phylogenetic tree analysis of MYB proteins in *Rhododendron lapponicum* and *Arabidopsis thaliana*

The transcription sequences of 75 002 *Rhododendron lapponicum* were compared with the NR protein database to obtain the annotation information of gene function. Among them, 71 155 transcripts had different degrees of homology with known genes of other species, accounting for 94.87% of the total number, indicating that SMRT sequencing technology is an effective method to excavate functional genes of *Rhododendron lapponicum*. In addition, 3 847 transcripts did not obtain functional annotation information, which may be due to the non-coding RNA sequence or short transcription sequence length, or the specific genes of *Rhododendron lapponicum* (Hou et al., 2011). GO database was used for functional classification of *Rhododendron lapponicum* transcripts, and 33 653 transcripts obtained specific functional classification entries. Among them, cell part, catalytic activity and cell process were the most transcripts in cell component, molecular function and biological process, respectively. Due to defects in the structural design of GO database and the fact that many features of genes have not been discovered, the functional classification information of GO gene is incomplete, and other bioinformatics methods are needed to supplement the functional annotations of transcripts (Jia et al., 2014). COG function classification was performed to further predict the function of genes expressed in *Rhododendron lapponicum*. Compared with COG database, 30 359 transcripts of *Rhododendron lapponicum* were divided into 24 COG function categories, and the proportion of general function prediction genes was the highest. In this study, the number of functional genes such as signal transduction mechanism, amino acid transport and metabolism, carbohydrate transport and metabolism, post-translational modification-protein turnover-molecular chaperone was large, indicating that genes such as metabolism, signal transduction, transcription and translation were relatively abundant in *Rhododendron lapponicum*.

As one of the largest gene families in plants, MYB transcription factor is involved in plant growth and development, signal transduction, substance metabolism, stress response, anthocyanin synthesis and other physiological and biochemical processes, but no relevant studies have been carried out in *Rhododendron lapponicum*. Compared with species without reference genome information, transcriptome sequencing data can be used to mine transcription factor families. For example, 117, 165 and 83 *MYB* genes can be identified from *Vaccinium Vitis-Idaea*, *Rehmannia glutinosa* (Gaetn.) Libosch. ex Fisch. et Mey., and *Lycium ruthenicum* Murr., etc. (Li et al., 2012; Wang et al., 2015; Yan Li et al., 2017). In this study, 64 transcription factor families, including 220 MYB family genes, were identified from *Rhododendron lapponicum* transcriptome data. The MYB transcription factor family can be divided into four categories based on the number of R structures contained in MYB transcription factors and the different types of repeats (Dubos et al., 2010). Studies have found that 1R-MYB and R2R3-MYB proteins are abundant in plants, while R1R2R3-MYB and 4R-MYB proteins are rare in plants (Niu et al., 2016). In this study, 157 1R-MYB proteins and 60 R2R3-MYB proteins were identified, while only 1 and 2 R1R2R3-MYB and 4R-MYB proteins were identified. Phylogenetic tree analysis showed that some *Rhododendron lapponicum* MYB subsets did not include Arabidopsis *MYB* genes, and some Arabidopsis MYB subsets did not include *Rhododendron lapponicum MYB* genes, which was similar to the results of *Rhododendron delavayi*. In this study, SMRT sequencing technology was used to sequence the transcriptome of *Rhododendron lapponicum*. The obtained transcripts were functionally annotated and classified, and 220 MYB transcription factor genes were identified, which laid a foundation for studying the gene structure and biological function of the MYB transcription factor family.

3 Materials and Methods

3.1 Test materials

The *Rhododendron lapponicum* 'Fuli Jinling' was collected from the Germplasm Resource Nursery of Rhododendron of Jiangsu Academy of Agricultural Sciences. In April 2019, the roots, stems, leaves, flowers and other tissue samples of "Fufi Jinling" during the flowering period were collected and put into ziplocked bags, which were immediately frozen with liquid nitrogen and stored in a refrigerator at -70°C for later use.

3.2 Library construction and sequencing

Total RNA was extracted from different tissues (roots, stems, leaves and flowers) by kit, and the integrity, purity and quality of total RNA were detected by agarose gel electrophoresis, Agilent2100 analyzer and NanoDrop2000 spectrophotometer. RNA samples from 4 different tissues were mixed in equal quantities to construct sequencing libraries. Transcriptome determination and assembly analysis were entrusted to Biomarker Technologies. The samples were sequenced on PacBio RS II platform and the library was established to obtain the original polymerase reading sequence. After reading the sequence and removing the joint, SMRT analysis software was used to identify, classify, cluster and correct the inserted fragment, finally obtaining high-quality full-length consistent sequence. The high-quality full-length sequences from the library were merged together, and CD-HIT was used to remove redundancy from the clustered and error-corrected transcripts.

3.3 Functional annotation and classification of *Rhododendron lapponicum* genes

BLAST software was used to compare the non-redundant transcripts with the NR database to obtain the functional annotation information of the transcripts. According to the functional annotation information of NR database, Blast2GO software was used to obtain GO entries of transcripts, and then WEGO software was used to conduct GO function classification statistics of transcripts (Hu et al., 2017). The transcripts were compared with COG database to obtain COG functional annotation and classification.

3.4 Prediction of *Rhododendron lapponicum* transcription factor family

Transcription sequences from *Rhododendron lapponicum* transcriptome data were submitted to the online database CD-HIT Suite to remove redundant sequence fragments, and then the redundant transcription sequences were submitted to the plant transcription factor database PlantTFDB for transcription factor prediction.

3.5 Classification and physicochemical properties analysis of MYB transcription factors in *Rhododendron lapponicum*

The MYB transcription factor was selected from the predicted transcription factor database of *Rhododendron lapponicum*. The ORF Finder online software of NCBI was used to analyze and predict the open reading frame of each *MYB* gene, and Smart BLAST was compared with the corresponding protein database for identification. ExPASy-pROSITE online software was used to predict and classify the number of DNA-binding domains contained in the N-terminus of MYB transcription factor.

3.6 Conserved elements and phylogenetic tree analysis of MYB transcription factor in *Rhododendron lapponicum*

The conserved elements of MYB gene in *Rhododendron lapponicum* were analyzed by MEME software. The MYB transcription factor of *Arabidopsis thaliana* was compared with the predicted MYB transcription factor of *Rhododendron lapponicum* by multi-sequence alignment function of MEGA7.0 software Clustal W, and phylogenetic tree was constructed.

Authors' Contributions

JXP is the executor of this experiment, responsible for experimental design, implementation, data analysis and writing the first draft of the paper. LXQ, DYM and LC participated in data analysis, formation and modification of the first draft; SJL determined the conception of the research project and directed the writing and revision of the paper. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (31700627) and the Natural Science Foundation of Jiangsu Province (BK20170607).

References

Ardui S., Ameur A., Vermeesch J.R., and Hestand M.S., 2018, Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics, Nucleic Acids Research, 46(5): 2159-2168
https://doi.org/10.1093/nar/gky066
PMid:29401301 PMCid:PMC5861413

Branton D., Deamer D.W., Marziali A. and Bayley H., 2008, The Potential and Challenges of Nanopore Sequencing, Nature Biotechnology, 26(10): 1146-1153
https://doi.org/10.1038/nbt.1495
PMid:18846088 PMCid:PMC2683588

Chen T., Li S.F., Qu S.P., Qian X., Liu J., and Xie W.J., Growth adaptability of 18 Rhododendron lapponicum cultivars in central Yunnan region, Xinan Nongxue Xuebao (Southwest China Journal of Agricultural Sciences), 32(3): 609-614

Dai X.Y., Xu Y.Y., Ma Q.B., Xu W.Y., Wang T., Xue Y.B., and Chong K., 2007, Overexpression of an R1R2R3 MYB Gene, OsMYB3R-2, increases tolerance to freezing, drought, and salt stress in transgenic Arabidopsis, Plant Physiol., 143(4): 1739-1751
https://doi.org/10.1104/pp.106.094532
PMid:17293435 PMCid:PMC1851822

Dubos C., Stracke R., Grotewold E., Weisshaar B., Martin C., and Lepiniec L., 2010, MYB transcription factors in Arabidopsis, Cell, 15(10): 573-581
https://doi.org/10.1016/j.tplants.2010.06.005
PMid:20674465

Hobert O., 2008, Gene regulation by transcription factors and microRNAs, Science, 319(5871): 1785-1786
https://doi.org/10.1126/science.1151651
PMid:18369135

Hou R., Bao Z.M., and Wang S., 2011, Transcriptome sequencing and de novo analysis for yesso scallop (Patinopecten yessoensis) using 454 GS FLX, PLoS One, 6(6): 118
https://doi.org/10.1371/journal.pone.0021560
PMid:21720557 PMCid:PMC3123371

Hu J.J., Meng X., Zhou J.B., Yang L.X., Liu S.H., and Li R.Z., 2017, Transcriptome analysis of the cotton mealybug, *Phenacoccus solenopsis* (Hemiptera: Pseudococcidae), Kunchong Xuebao (Acta Entomologica Sinica), 60(1): 9-17

Huang C.C., Hung K.H., Hwang C.C., Huang J.C., and Chiang T.Y., 2011, Genetic population structure of the alpine species *Rhododendron pseudochrysanthum* sensu lato (Ericaceae) inferred from chloroplast and nuclear DNA, BMC Evolutionary Biology, 11(1): 108
https://doi.org/10.1186/1471-2148-11-108
PMid:21501530 PMCid:PMC3096940

Jia X., Deng Y., Sun X., Liang L., and Su J., 2016, De novo assembly of the transcriptome of Neottopteris nidus using illumina paired-end sequencing and development of EST-SSR markers, Molecular Breeding, 36(7): 94
https://doi.org/10.1007/s11032-016-0519-2

Jia X., Tang L., Mei X., Liu H., and Su J., 2020, Single-molecule long-read sequencing of the full-length transcriptome of Rhododendron lapponicum, Scientific Reports, 10(1): 6755
https://doi.org/10.1038/s41598-020-63814-x
PMid:32317724 PMCid:PMC7174332

Jia X.P., Sun X.B., Deng Y.M., Liang L.J., and Ye X.Q., 2014, Sequencing and analysis of the transcriptome of Asplenium nidus, Yuanyi Xuebao (Acta Horticulturae Sinica), 41(11): 2329-2341

Kranz H.D., Denekamp M., Greco R., Jin H., Leyva A., and Meissner R.C., 2010, Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*, Plant Journal, 16(2): 263-276
https://doi.org/10.1046/j.1365-313x.1998.00278.x
PMid:9839469

Li C.Q., Wang Y., Huang X.M., Li J., Wang H.C., and Li J.G., 2013, De novo assembly and characterization of fruit transcriptome in Litchi chinensis Sonn and analysis of differentially regulated genes in fruit in response to shading, BMC Genom., 14: 552
https://doi.org/10.1186/1471-2164-14-552
PMid:23941440 PMCid:PMC3751308

Li X., Sun H., Pei J., Dong Y., Wang F., and Chen H., 2012, De novo sequencing and comparative analysis of the blueberry transcriptome to discover putative genes related to antioxidants, Gene, 511(1): 54-61
https://doi.org/10.1016/j.gene.2012.09.021
PMid:22995346

Li X.L., Luo L.L., and Hua R.Z., 2018, Physiological and biochemical responses of Rhododendron lapponicum to heat stress, Xibei Zhiwu Xuebao (Acta Botanica Boreali-occidentalia Sinica), 27(2): 253-259

Liu X., Mei W., Soltis P.S., Soltis D.E., and Barbazuk W.B., 2017, Detecting alternatively spliced transcript isoforms from single-molecule long-read sequences without a reference genome, Molecular Ecology Resources, 17(6): 1243-1256
https://doi.org/10.1111/1755-0998.12670
PMid:28316149

Liu X.Q., Su J.L., Li C., He L.S., and Chen L., 2011, A new Alpine Rhododendron cultivar'Fuli Jinling', Yuanyi Xuebao (Acta Horticulturae Sinica), 38(11): 2237-2238

Niu Y.L., Jiang X.M., and Xu X.Y., 2016, Reaserch advances on transcription factor MYB gene family in plant, Fenzi Zhiwu Yuzhong (Molecular Plant Breeding), 14(8): 2050-2059

Wang F., Suo Y., Wei H., Li M., Xie C., and Wang L., 2015, Identification and characterization of 40 isolated Rehmannia glutinosa MYB family genes and their expression profiles in response to shading and continuous cropping, International Journal of Molecular Sciences, 16(7): 15009-15030
https://doi.org/10.3390/ijms160715009
PMid:26147429 PMCid:PMC4519885

Wang Z.Y., Fang B.P., Chen J.Y., Zhang X.J., Luo Z.X., Huang L., Chen X., and Li Y., 2010, De novo assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweet potato (Ipomoea batatas), BMC Genom., 11: 726
https://doi.org/10.1186/1471-2164-11-726
PMid:21182800 PMCid:PMC3016421

Wei W.L., Qi X.Q., Wang L.H., Zhang Y.X., Hua W., Li D., Lv H., and Zhang X., 2011., Characterization of the sesame (Sesamum indicum L.) global transcriptome using Illumina paired-end sequencing and development of EST–SSR markers, BMC Genom., 12: 726
https://doi.org/10.1186/1471-2164-11-726
PMid:21182800 PMCid:PMC3016421

Xiao J., Jin X.H., Jia X.P., Wang H.Y., Cao A.Z., Zhao W.P., Pei H.Y., Xue Z.K., He L.Q., Chen Q.G., and Wang X.E., 2013, Transcriptome-based discovery of pathways and genes related to resistance against Fusarium head blight in wheat landrace Wangshuibai, BMC Genom., 14: 197
https://doi.org/10.1186/1471-2164-14-197
PMid:23514540 PMCid:PMC3616903

Yan L., Wang C.P, Chen J.W., Qiao G.X., and Li J., 2017, Analysis of MYB transcription factor family based on transcriptome sequencing in Lycium ruthenicum Murr., Zhongguo Nongye Kexue (Scientia Agricultura Sinica), 50(20): 3991-4002

Zhang L.W., Wan X.B., Xu J.T., Lin L.H., and Qi J.M., 2015, De novo assembly of kenaf (Hibiscus cannabinus) transcriptome using Illumina sequencing for gene discovery and marker identifification, Mol. Breeding, 35: 192
https://doi.org/10.1007/s11032-015-0388-0