

R2R3-MYB transcription factors from *Lotus corniculatus* var. *japonicus* differentially regulated in roots and shoots in response to nitrogen starvation.

KUNIIHIKO MIYAKE¹, TAKURO ITO¹, MINEO SENDA², RYUJI ISHIKAWA¹, TAKEO HARADA¹,
MINORU NIIZEKI¹ and [SHINJI AKADA](#)^{2*}

¹ Plant Breeding Laboratory, Faculty of Agriculture, Hirosaki University, Hirosaki, Aomori
036-8561, Japan

² Gene Research Center, Hirosaki University, Hirosaki, Aomori 036-8561, Japan

* Corresponding author

In leguminous plants, a species-specific type of (iso)flavonoid secreted out from roots functions as a trigger for the nitrogen-fixing symbiosis of rhizobacteria. The pathway of (iso)flavonoid biosynthesis is significantly down-regulated by exogenously supplied nitrate in alfalfa (Coronado *et al.*, 1995). This suggested a possibility that inhibitory effects of nitrate on nodulation, in part, operate via regulation of the pathway of flavonoid biosynthesis. On the other hand, it has been reported that pathways of flavonoid biosynthesis are regulated by various members of *R2R3-MYB* superfamily in different organs and in different conditions (Paz-Ares *et al.*, 1987; Grotewold *et al.*, 1994; Quattrocchio *et al.*, 1999; Jin *et al.*, 2000; Nesi *et al.*, 2001). In this study, we attempted to isolate members of the *R2R3-MYB* superfamily, which may be involved in regulation of nitrogen fixation or of nitrogen metabolism in legumes.

A library of 300 recombinant plasmid clones containing the PCR fragments amplified by using degenerate primers for the *R2R3-MYB* superfamily was screened by differential hybridization to isolate *R2R3-MYB* genes whose expression was up-regulated under nitrogen nutrient limited conditions. Two groups of clones were identified, each of which seemed to represent a gene responsive to the nitrogen starvation. The entire coding regions for the genes were further isolated by PCR and were designated as *LjMYB101* and *LjMYB102*. By screening a genomic library of *Lotus japonicus* with a probe derived from *LjMYB101*, the third gene, *LjMYB103*, was isolated. In addition, a candidate for the soybean orthologue of *LjMYB101* was isolated and designated as *GmMYB101* (Miyake *et al.*, 2003). Sequence alignment of the genes with members of the plant *R2R3-MYB* superfamily showed that they all belonged to the subgroup 10 (Stracke *et al.*, 2001) of the *R2R3-MYB* superfamily (Figure 1). Among these genes a contrasting expression pattern has been observed between *LjMYB101* and *LjMYB103*. That is, the expression level of *LjMYB101* was higher in roots and it was up-regulated under a nitrate limited condition (N-) in the roots, but not in the shoots. On the other hand, expression of *LjMYB103* was higher in shoots and was also up-regulated only in the shoot (Figure 2). Considering their structural similarity in the coding region, *LjMYB101* and *LjMYB103* seem to have diverged through acquisition of their

organ specific responsiveness, which might have differentiated their physiological roles. Therefore, the primary role of *LjMYB103* might have been assigned to the shoot while the same kind of role fulfilled by *LjMYB101* in the roots.

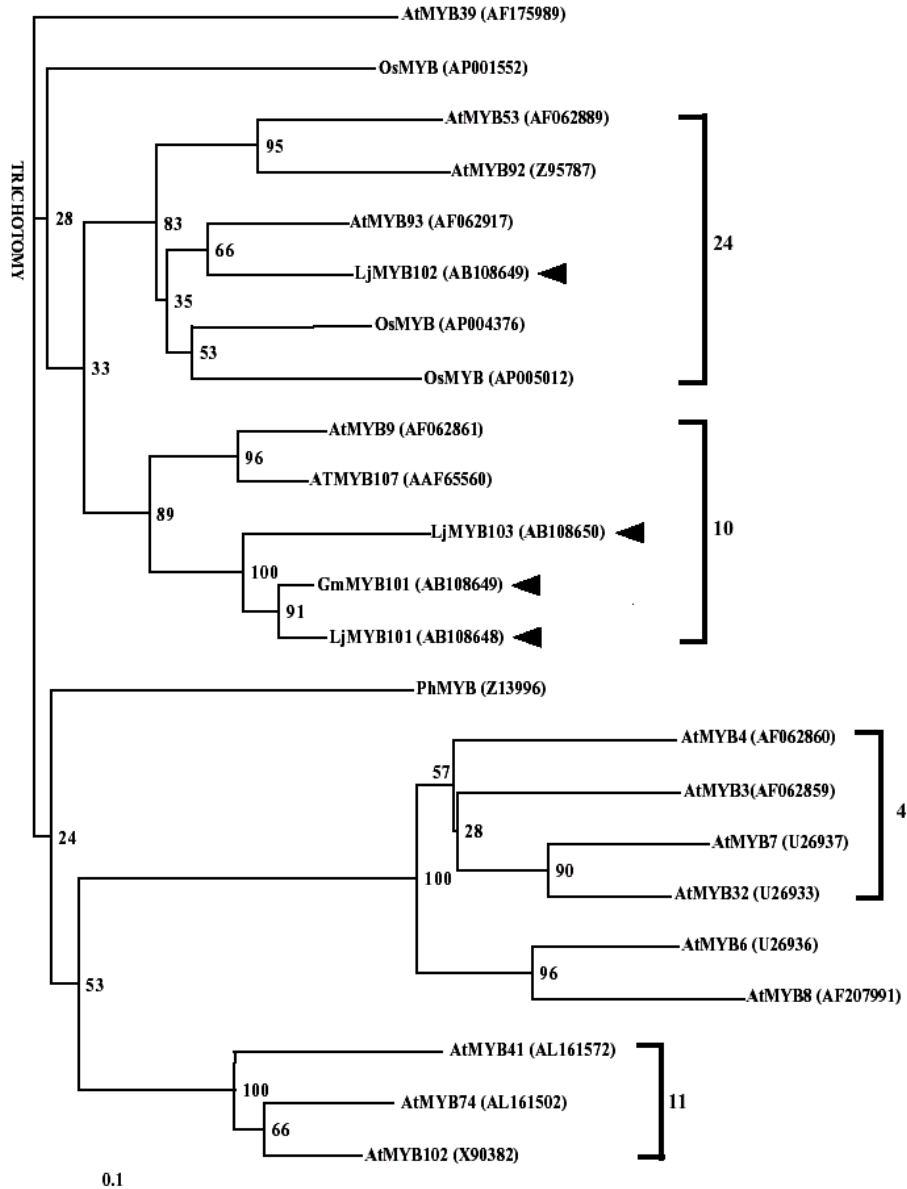


Figure 1. Phylogeny of the members in the subgroup 10 as well as those in its related subgroups (24, 4, and 11) of the plant MYB superfamily. The amino acid sequences of *LjMYB101*, *LjMYB102*, *LjMYB103*, and *GmMYB101* (indicated with arrowheads) along with those of the genes of which sequences are currently in GenBank/EMBL/DBJ (accession numbers are indicated in parenthesis) were aligned with the ClustalW (Thompson *et al.*, 1994) and, at the same time, reconstruction of a phylogenetic tree with neighbor joining method and calculation the bootstrap values were carried out with the program package of DBJ. Members of each subgroup are indicated with side brackets.

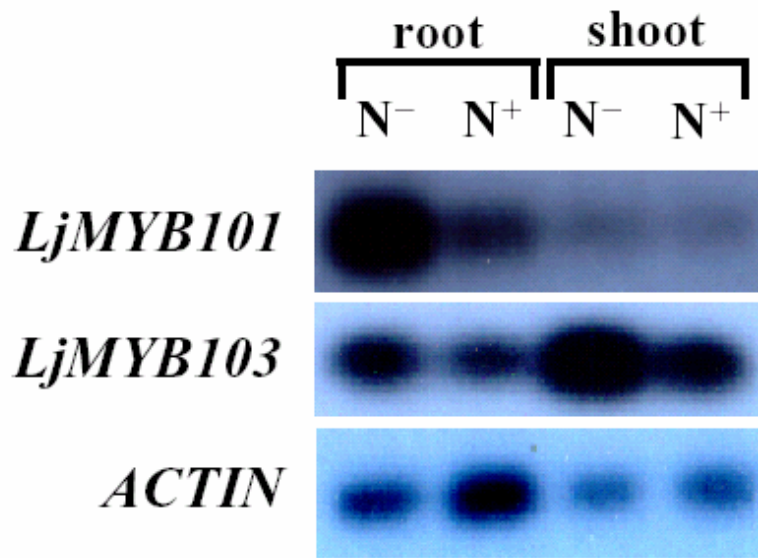


Figure 2. Differential expression of *LjMYB101* and *LjMYB103* in the roots and shoots of *L. japonicus* grown under the N⁻ or N⁺ conditions. RT-PCR/Southern analysis was carried out using the specific primers for each gene. The total RNA samples as templates were extracted from seedlings grown with the agar medium containing the basic MS salts (Murashige and Skoog, 1962) with its nitrogenous components (NH₄NO₃ plus KNO₃) substituted for by the equal molar (39.4 mM) of KCl (N⁻) or KNO₃ (N⁺). The expressed levels of *LjACTIN* transcripts were analyzed as the reaction controls.

At this point, we don't have any evidence defining the cis-elements that conferred their organ specificity, but the sequence information of the promoters seems informative. The 5' upstream fragment (524 bp) of *LjMYB101* was isolated by inversed PCR and that (1569 bp) of *LjMYB103* was obtained from the lambda genomic clone, λ Lj72302-2. On the analysis of the sequence information, a striking similarity was found between the promoter segments of *LjMYB101* and *LjGln1*, a member of the Gln synthetase gene family in *Lotus corniculatus* var. *japonicus* (ThykJær *et al.*, 1997; Figure 3). The homologous region spans over a stretch of 138 bp with similarity of 89 %. It is located from 374 to 511 bp upstream of ATG in *LjMYB101*, or from 1517 to 1654 bp upstream in *LjGln1*, respectively. It would be interesting to speculate that insertion or deletion of this element in the promoter of the common ancestor of *LjMYB101* and *LjMYB103* might have caused diversion of the two genes. Because nitrate regulated expression of *LjGln1* was also restricted to the root tissues, the homologous region of *LjMYB101* and *LjGln1* may be related to their organ specific regulation. When a blast search was performed with the element of 138 bp, a list of 27 sequences containing this element was found in the genome of *Lotus corniculatus* var. *japonicus*. We found ORFs of various sizes in many of their proximal regions, but we did not find any genes closely relevant to nitrogen metabolism. Further work will be required to verify the significance of this element.

GLN1 ELEMENT	-1668	ATTTTAAAT	TTTATTTT	ATTCAAATAT	AAATITGATA	CATTTTTC	-1621
LJMYB101 PRO	-523	AATTTAAAT	GTATTTT	ATTCAAATAT	AAATCTGATA	CATTTTTC	-475
LJMYB103 PRO	-499	AACTT--AT	GT----CCCC	ATT-----T	AGGCT----	CATTTCCCAC	-466
GLN1 ELEMENT	-1620	TTATCATATC	ATGTCCTTAT	CATGTGCACC	TC-AAACATA	AGATAAGATA	-1572
LJMYB101 PRO	-474	TTATCATATC	CTCTCTTAT	CATGTGCACC	TC-AAACACA	GGATATGATA	-426
LJMYB103 PRO	-465	TAATTAAT	GTTCTTTA-	-A-GTATACC	TCTGAAT-TA	TAATTITGTG	-420
GLN1 ELEMENT	-1578	AAAGTCTATT	CTGCCCTTAT	GTATATCCTGT	TGCTATCATG	TTCACATATC	-1522
LJMYB101 PRO	-425	AGAGTCTATC	CTGCTTTTAT	CTATCCTGT	TGATATCCTG	TTCACATATC	-376
LJMYB103 PRO	-419	TGAG--TG--	--GTTTT--	--TAATAAGA	TGACA--CAA	TTTA-AT-TA	-384
GLN1 ELEMENT	-1528	ATATC-----	ATGTTCTA--	--ACCACCA	AACG-GACAT	A--G--AATT	-1487
LJMYB101 PRO	-375	ATATCCTATC	ATATCCTATC	CTGACCACCA	AACGAGCCCT	AA-G--AGTA	-329
LJMYB103 PRO	-383	ATAGAAAATG	AGACACAAA-	TGAGGA--G	AATGGATTTC	AATGCCAGTA	-338
LJMYB101 PRO	-328	GATAC--ATT	ACTGTGAGGA	GTGTCTTATG	ACCCCTTCA	AGGTCGTTA-	-282
LJMYB103 PRO	-337	AAAATTGGTA	ATT-TGAGGA	GTGTCTTGCC	ATTCTCTTGT	AGGTCGTTAC	-289
LJMYB101 PRO	-281	---TCTAC-	AT----GTA	TAATA-----	---CAGGACA	AATTTGGAAT	-250
LJMYB103 PRO	-288	TATCTCTATG	ACTTTGAGTG	TAATAGACAT	AAACAGGACA	AAATTGAAAT	-239
LJMYB101 PRO	-249	TTTATATGA-	GA-ACATGAT	---AAA--C	A-AG-T----	CTTATATTTC-	-215
LJMYB103 PRO	-238	TTGCCTTGAC	AACACATGAT	GTACAAAGTC	AGACATGGCA	CTGATATTTT	-189
LJMYB101 PRO	-214	CAAACATGTC	CATATGCTAT	ATATAACAAG	CACCCCTCTA	CTAAATTAACC	-165
LJMYB103 PRO	-188	CAAACATGAC	--TACCTTAT	ATATAA-AA-	----TCCCTG	GTAATTAAGTT	-147
LJMYB101 PRO	-164	CAAGTCCCC	CACAGTACTC	TAATCTATAT	ACCTTCATTC	TCCAGTTTCC	-115
LJMYB103 PRO	-146	---GTT---	----GTCCTC	CACCTCACAC	A-----A---	---CAGTTCA-	-118
LJMYB101 PRO	-114	AACACTCCAA	GGTATGGATT	GATT-TGCT-	---CTCTTTC	ATTGCCAT-G	-71
LJMYB103 PRO	-117	---ATTG-AA	GGTTGGATT	G-TTGTGCTG	GTTCTCTTTC	TTTCTTTTC	-73
LJMYB101 PRO	-70	ACCATGAAAG	TTATCC-CTA	GAATCAAAC	AAAC-TTGT	ACTATTTTGA	-23
LJMYB103 PRO	-72	GTCCATGTAA	CTACTAACT-	GAATCAAAC	AA-CTATTGT	ACTAT-----	-30
LJMYB101 PRO	-22	ATGTGTTTGT	GCAG-TGCA	GTGATG.....	+3
LJMYB103 PRO	-29	ATAT-TTCTT	CTAGATACAT	AT-ATCAGT	ATG.....	+3

Figure 3. Sequence alignment of the promoters of *LjMYB101* (LjMYB101 PRO) and *LjMYB103* (LjMYB103 PRO), as well as the homologous sequence element found in the promoter of LjGln1 (GLN1 ELEMENT). The sequences were aligned using the Higgins' multiple alignment program of DNASIS V3.6 (HITACH SOFTWARE). The TATA-boxes are underlined and the homologous elements of 138 bp are lined above the sequences.

Downstream of this promoter element is a promoter core region including 332 bp upstream of the ATG initiation codon in *LjMYB101* or 341 bp upstream of *LjMYB103*, respectively. The putative TATA-box was found at 190 bp or 166 bp upstream of ATG for each of the genes. Sequence alignment of the two promoters revealed that *LjMYB101* and *LjMYB103* shared continually homologous sequences around the regions surrounding the TATA-box (Figure 3). Thus, the two sequences seem to have retained substantial homology, while little homology was detected between the two in their further upstream regions, because of presence of the element of 138 bp in the promoter of *LjMYB101*. Albeit their difference in tissue specificity, the two genes are both up-regulated under nitrogen-limited conditions. Thus, it is possible that the sequence element conserved between the promoters of *LjMYB101* and *LjMYB103* might act as *cis*-elements for the nitrogen regulation.

We speculated that LjMYB101 might function as a transcriptional activator of genes such as members of chalcone synthase gene family (*LjCHS*) and/or *LjGln1*, because both these genes are also up-regulated specifically in roots under nitrogen-limited conditions (Miyake *et al.*, 2003). In order to verify if the MYB transcription factor targets to the promoters of these genes, the DNA binding domain of LjMYB101 was synthesized using an *E.coli* protein expression system. The protein was used for the electrophoretic mobility-shift assay (EMSA) and it was found that the protein bound to the probe DNA fragments containing the sequence of CACCTACC or CACCAACC, which coincided very well with the type IIG (H-box like) consensus of YACCWACC as a recognition sequence for a class of MYB transcription factors. The sequence perfectly matching YACCWACC is found in three members (accession numbers; AP006705, AP006706, and AP006709) of *LjCHS*, among eight members of the gene family available from the database. This motif locates upstream of the TATA-box and the distance ranges from 22 to 34 bp, which is very similar to the distance of the same element to the TATA-box on the promoter regions of the soybean *CHS* gene family, ranging from 22 to 38 bp (Shimizu *et al.*, 1999). These suggest that LjMYB101 may target to the promoters of members of *CHS* gene family in *Lotus* while GmMYB101 may do the same in soybean. On the other hand, the sequence of CACCTACC also exists at 43 bp upstream of TATA-box on the promoter of *LjGln1*, inspiring another possibility that LjMYB101 may target to *LjGln1*. In vivo analysis of the gene function will be required to further investigate the physiological roles for the newly isolated R2R3-MYB genes, in relation to flavonoid biosynthesis or to nitrogen metabolism.

References.

- CORONADO C., ZUANAZZI J.A.S., SALLAUD C., QUIRION J.-C., ESNAULT R., HUSSON H.-P., KONDOROSI A. and RATET P. 1995. Alfalfa root flavonoid production is nitrogen regulated. *Plant Physiology*, **108**, 533-542.
- GROTEWOLD E., DRUMMOND B.J., BOWEN B. and PETERSON T. 1994. The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell*, **76**, 543-553.
- HITACH SOFTWARE. Hitachi Software Engineering Co., Ltd.

- JIN, H., COMINELLI, E., BAILEY, P., PARR, A., MEHRTENS, F., JONES, J., TONELLI, C., WEISSHAAR, B., AND MARTIN, C. 2000. Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *The EMBO Journal*, **19**, 6150-61.
- MIYAKE K., ITO T., SENDA M., ISHIKAWA R., HARADA T., NIIZEKI M. and AKADA S. 2003. Isolation of a subfamily of genes for R2R3-MYB transcription factors showing up-regulated expression under nitrogen nutrient-limited conditions. *Plant Molecular Biology*, **53**, 237-245.
- MURASHIGE T. and SKOOG F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum*, **15**, 473-497.
- NESI N., JOND C., DEBEAUJON I., CABOCHE M. and LEPINIEC L. 2001. The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *The Plant Cell*, **13**, 2099-114.
- PAZ-ARES J., GHOSAL D., WIENAND U., PETERSON P.A. and SAEDLER H. 1987. The regulatory *c1* locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *The EMBO Journal*, **6**, 3553-3558.
- QUATTROCCHIO F., WING J., WOUDE K., SOUER E., VETTEN N., MOL J. and KOES R. 1999. Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. *The Plant Cell*, **11**, 1433-1444.
- SHIMIZU T., AKADA S., SENDA M., ISHIKAWA R., HARADA T., NIIZEKI M. and DUBE S.K. 1999. Enhanced expression and differential inducibility of soybean chalcone synthase genes by supplemental UV-B in dark-grown seedlings. *Plant Molecular Biology*, **39**, 785-795.
- STRACKE R., WERBER M. and WEISSHAAR B. 2001. The R2R3-MYB gene family in Arabidopsis thaliana. *Current Opinion in Plant Biology*, **4**, 447-56.
- THOMPSON L.D., HIGGINS D.G. and GIBSON T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.
- THYKJÆR T., DANIELSEN D., SHE Q. and STOUGAARD J. 1997. Organization and expression of genes in the genomic region surrounding the glutamine synthetase gene Gln1 from *Lotus japonicus*. *Molecular Genetics and Genomics*, **255**, 628-636.