

An extensive analysis of *R2R3-MYB* regulatory genes from *Fagus crenata*

Shuichi Matsuda · Naoko Wakamatsu · Hitomi Jouyu · Hajime Makita · Shinji Akada

Received: 25 November 2009 / Revised: 2 August 2010 / Accepted: 25 August 2010 / Published online: 27 October 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Using pairs of degenerate primers, we conducted a polymerase chain reaction to amplify the partial R2R3 domains of a majority of the *R2R3-MYB* family genes from *Fagus crenata* and identified a total of 85 independent gene fragments. By phylogenetic analysis of the deduced amino acid sequences, we found that many of the beech genes clustered with members from *Arabidopsis*, suggesting that these members represent beech orthologs of *Arabidopsis*. Some of the orthologous relationships became more evident when the complete gene structures were compared. Further, a large number of genes formed an additional and expanding cluster, independent from the other subgroups. These members were further compared with the *Populus* and *Vitis* family genes. In the epidermal cell fate clade, expansion of the beech family genes was comparable with those of the *Populus* and *Vitis* families, but the number of genes present in every subclade fluctuated extensively.

Communicated by M. Byrne

Electronic supplementary material The online version of this article (doi:10.1007/s11295-010-0333-7) contains supplementary material, which is available to authorized users.

S. Matsuda
The United Graduate School of Agricultural Sciences,
Iwate University,
3-18-8 Ueda,
Morioka, Iwate 020-8550, Japan

N. Wakamatsu · H. Jouyu · H. Makita
Faculty of Agriculture and Life Science, Hirosaki University,
3 Bunkyo-cho,
Hirosaki, Aomori 036-8561, Japan

S. Akada (✉)
Gene Research Center, Hirosaki University,
3 Bunkyo-cho,
Hirosaki, Aomori 036-8561, Japan
e-mail: akada@cc.hirosaki-u.ac.jp

Beech genes were abundant in the general flavonoid pathway regulation and TT2-related subclades; no beech gene was included in the anthocyanin-related subclade. Further analysis of the newly amplified regulatory genes to elucidate their functions may clarify the role of these genes in the evolution of plant species.

Keyword R2R3-MYB · Beech · Gene family · Environmental stresses

Introduction

The genetic analysis of samples collected from Japanese beech forests has made it increasingly more apparent that these forests are extremely diverse (Tomaru et al. 1997; Takahashi et al. 2000; Asuka et al. 2004). This finding implies that various mutations have gradually accumulated in the forest population throughout the long Quaternary history of their vegetation in Japan. It is generally accepted that the genetic complexity of the forest tree species such as those in Fagaceae is maintained in part by means of long-distance dispersal of pollen and gravity-based distribution of seeds (Comps et al. 2001; Asuka et al. 2005). Heterosis and inbreeding depression may also have contributed to the maintenance of genetic diversity. However, in spite of the presence of polymorphic alleles in the population, we are still far from understanding what type of genes have actually contributed to the prolonged vegetation under the ever-changing environment. A possible method of identifying these genes may be the analysis of genes that function in response to environmental stresses and those that are required for adaptation to various living conditions.

In plants, R2R3-MYB transcription factors form a large family of around 100 members. They all share a conserva-

tive R2R3-DNA-binding domain (R2R3-DBD) in their N-terminal regions, whereas their C-terminal parts are highly variable (Paz-Ares et al. 1987; Romero et al. 1998; Kranz et al. 1998; Shimizu et al. 2000). In *Arabidopsis thaliana*, the *R2R3-MYB* gene family is divided into 24 subgroups, with each subgroup distinguished by some specific motifs in the C-terminal part (Kranz et al. 1998; Stracke et al. 2001). A variety of physiological roles have been described for different members of the family. Many of these genes participate in the modulation of biological processes such as epidermal cell patterning, vegetative growth and development, dark photomorphogenesis, pollen development, tapetum development, petal formation, hypersensitive cell death, and stomatal movements (Oppenheimer et al. 1991; Kirik et al. 1998; Vailleau et al. 2002; Higginson et al. 2003; Newman et al. 2004; Preston et al. 2004; Cominelli et al. 2005; Baumann et al. 2007). The direct targets of these genes are often genes involved in the pathways of phenylpropanoid, tryptophan, or fatty acid biosynthesis (Bender and Fink 1998; Preston et al. 2004; Stracke et al. 2007; Park et al. 2008; Raffael et al. 2008). Some of these genes are involved in signaling pathways that are activated to cope with environmental stresses such as cold weather, drought, ultraviolet (UV)-B irradiation, low oxygen levels, and rhizobacteria-mediated induced systemic resistance (Urao et al. 1996; Hoeren et al. 1998; Jin et al. 2000; Zhu et al. 2005; Van der Ent et al. 2008). Thus, the majority of the *R2R3-MYB* family members seem to be involved in the elaboration of the well-adjusted structure and the fine-tuned function of plant tissues and organs in response to the changing environment. Here, we focused on the isolation and analysis of genes belonging to the *R2R3-MYB* family from beech trees in order to determine if any of these genes are involved in environmental adaptation.

Using pairs of degenerate primers, we polymerase chain reaction (PCR)-amplified most of the *MYB* family genes from *Fagus crenata* and identified a total of 85 independent genes. On the basis of the sequence information of these genes, we conducted a comprehensive analysis of the regulatory *R2R3-MYB* genes of the beech. We discuss how each gene of the beech *MYB* family could have developed to date in its own way and what specific genes of the beech could tell us in relation to the genetic diversity of the species.

Materials and methods

Plant materials and methods of DNA and RNA isolation

Leaves or twigs with winter buds were collected from beech trees on Mt. Shirakamidake (at an altitude of 400–

1,150 m) and Mt. Masedake (at an altitude of 300–1,100 m) and numbered S400-1–5, S600-1–5, S800-1–5, S1000-1–5, and S1150-1–5 or M300-1–5, M500-1–5, M700-1–5, M900-1–5, and M1000-1–5. Twigs with winter buds were incubated in a greenhouse at 25°C until young leaves emerged from the buds. Total DNA was isolated from the leaf tissues by using the cetyltrimethylammonium bromide method as described by Richards et al. (1994).

Total RNA from mature leaves collected in the fall (beginning of October) from a beech tree growing on the campus of Hirosaki University was isolated using the guanidium thiocyanate/CsCl method as described before (Shimizu et al. 1999). RNA samples of young leaves, petioles, male flowers, and pollen collected in the end of April from the same tree on the campus were isolated using an RNA-extraction kit (QIAGEN RNeasy Plant Mini Kit) according to the product instructions.

For isolation of stress-responsive genes, twigs with five or six winter buds were sampled from the beech tree on the campus and incubated with their lower ends dipped in water at 25°C under a 16/8-h light/dark (LD) cycle until young leaves emerged from the buds. Subsequently, the 3-day-old young leaves were incubated under different stress conditions as follows: cold treatment—twigs were exposed to low (4°C) temperature for 3 or 6 h; high-temperature treatment—twigs were exposed to high (42°C) temperature for 30 min, and the young leaves were sampled immediately and at 1.5 h after they had been cooled to room temperature; drought condition—twigs were incubated at 25°C without dipping them into water for 3, 6, and 12 h; gibberellic acid treatment—lower ends of twigs were dipped in 0.1 mM of gibberellic acid solution at 25°C under a 16/8 LD cycle. Total RNA was isolated from these leaf tissues as described by Chang et al. (1993).

PCR amplification of a region of MYB encoding a 42-amino-acid-long product and genomic screening

A partial R2R3 domain of MYB was amplified using PCR or reverse transcription (RT)-PCR with degenerate primers (Romero et al. 1998; Shimizu et al. 2000) whose target sequences are depicted in Fig. S1. This domain encodes a 42-amino-acid-long conservative region. The amplification products were cloned into a T-vector prepared by the addition of dT to the *EcoRV*-digested blunt ends of pBluescript II SK⁺ (Stratagene), and the nucleotide sequence of each of the cloned DNA fragments was determined using automated DNA sequencers (LICOR or ABI) according to the manufacturers' instructions.

A beech genomic library was constructed with the total DNA isolated from S600-3 using a Lambda FIXII/*Xho*I Partial Fill-In Vector Kit (Stratagene) in accordance with the product instructions. A total of approximately 1.2×10^5

independent plaques were screened using as probes PCR fragments containing the MYB partial R2R3 domains of *FcMYB1901*, *FcMYB3202*, *FcMYB3103*, *FcMYB3201*, or *FcMYB2402*. The blots were prehybridized for 2 h at 42°C in a prehybridization buffer containing 50% formamide, 5×SSC, 5×Denhardtts, 25 mM Na phosphate buffer, and 0.2 mg/ml salmon sperm DNA. The blots were then hybridized overnight at 42°C after the addition of the ³²P-labeled probe to the hybridization buffer containing 50% formamide, 5×SSC, 1.8% SDS, 1×Denhardtts, 25 mM Na phosphate buffer, and 0.2 mg/ml salmon sperm DNA. The blots were washed twice in a solution of 2×SSC and 0.1% SDS for 5 min at 65°C, twice in a solution of 0.2×SSC and 0.1% SDS for 5 min at 65°C, and once in a solution of 0.2×SSC and 0.1% SDS for 5 min at 65°C. The hybridized blots were exposed to X-ray film (Fuji RX-U).

Phylogenetic analysis and protein sequence alignment of R2R3-MYB

All the identified genes, either genomic or cDNA sequences, were originally named after their sample sources and finally renamed systemically following the clustering results obtained by phylogenetic analysis described as below using the number of each cluster they belonged to along with the *Arabidopsis*, *Populus*, and *Vitis* R2R3-MYB family members. A phylogenetic tree and bootstrap values (5,000 replicates) based on the 42-amino-acid-long R2R3-DBD sequences of all the identified beech MYB products along with those of the corresponding regions of the products encoded by the members of the following R2R3-MYB families were generated using Jukes–Cantor evolutionary distances and the neighbor-joining method (Saitou and Nei 1987) in the MEGA 4 (Tamura et al. 2007) program. Branch length appears in the same units as those of the evolutionary distance used to infer the phylogenetic tree. Internal branch support was estimated with 5,000 bootstrap replicates. *Arabidopsis* (*A. thaliana*) R2R3-MYB gene identifiers were obtained from Stracke et al. (2001), and the corresponding protein sequences were downloaded from TAIR (<http://www.arabidopsis.org/>). *Populus trichocarpa* R2R3-MYB gene models were retrieved from the Joint Genome Institute *P. trichocarpa* version 1.1 website (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html). Gene identifiers for 36 *Vitis vinifera* R2R3-MYB genes in ‘epidermal cell fate’ clade were obtained from Matus et al. (2008), and the corresponding protein sequences were downloaded from the International Grape Genome Program’s (IGGP) website (http://www.genoscope.cns.fr/externe/English/Projets/Projet_ML/projet.html).

The PDF images of Clustal W (Thompson et al. 1994) output for the aligned full-length amino acid sequences were generated using Geneious v4.5.3. Amino acid residues

conserved among the compared sequences were indicated with white letters on gradient shades of background in the following criteria: black, 100% similar; dark gray, 80% to 100% similar; gray, 60% to 80% similar; white, less than 60% similar.

Quantitative RT-PCR analysis

Quantitative RT-PCR was carried out using the iScript cDNA synthesis kit (Bio-Rad) for the reverse transcriptase reactions and iQ Supermix (Bio-Rad) and the following primers and probes for the real-time PCR. For the target gene (*FcMYB1901*): sense primer, TGC ACC CAC AAC TCC ATC AAC; anti-sense primer, GGG TGA GAG TGG AAA TGG AAG G; probe, FAM-CCCAACCACACCAAG CACACCCAC-BHQ1; and for the housekeeping control gene (*FcACT*): sense primer, ATT CTC ACC GAG AGA GGT TAC ATG; anti-sense primer, AGT CTC AAG TTC CTG CTC ATA GTC; probe, HEX-TCACCACCACT GCCGAACGGGAAA-BHQ1. The reaction was monitored and analyzed with the Opticon real-time PCR system (Bio-Rad).

Nucleotide sequence data

The nucleotide sequence data reported here are available from the DDBJ/EMBL/GenBank databases under the accession numbers AB434546 to AB434652, AB499731 to AB499735, AB500067 to AB500085, AB558901, and AB558902.

Results

Random amplification of MYB partial R2R3 domain from total DNA

We first used a pair of primers, MYBdgFP1 and MYBdgRP1, for the PCR amplification of the MYB partial R2R3 domain from total DNA, and the PCR products from nine different individuals were analyzed using polyacrylamide gel electrophoresis (Fig. 1). Because the amplified region contained an intron between two exons, various sizes of fragments were detected from each individual, indicating that multiple genomic regions including the MYB partial R2R3 domain could be amplified using the degenerate primers. In addition, some fragments showed individual differences. For instance, the lowest band (size, approximately 250 bp) found in lanes 1–5 and 9 was not detected in lanes 6–8. In contrast, the second lowest band (size, approximately 260 bp) found in lanes 2 and 5–9 was not detected in lanes 1, 3, and 4. These individual differences suggested the presence of some polymorphic

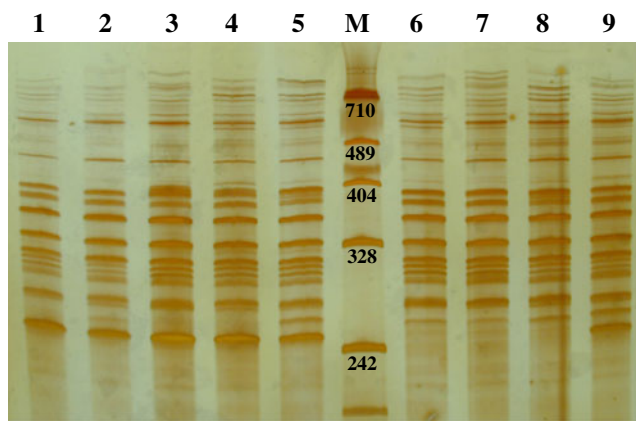


Fig. 1 PCR amplification of MYB partial R2R3-DBD from total DNA. The PCR products amplified with the primers MYBdgFP1 and MYBdgRP1 from the total DNA samples of the individuals numbered as 1, S400-5; 2, S600-5; 3, S800-5; 4, S1000-5; 5, S1150-5; 6, M300-5; 7, M700-5; 8, M900-5; 9, M1000-5 were separated by 12% polyacrylamide gel electrophoresis and detected with silver staining. *M*, *Msp*I-digested pBluescript II SK⁺ was used as a molecular size marker

genes, and the two smallest bands might represent alleles of a polymorphic *MYB* gene. The above supposition was subsequently proved by the sequence determination of the lowest band, which was then designated *FcMYB3202*.

The PCR products from S1150-5 (Fig. 1, lane 5) were shotgun-cloned into the bacterial plasmid vector pBlue-script II SK⁺, and after sequence determination 19 independent homologous *MYB* sequences were identified from the recombinant clones. Typically, the identified region contained two exon parts (92 bp from exon 2 and 35 bp from exon 3) and an intron sequence of various sizes between the two exons. The identified *MYB* sequences were originally named after the total sizes of the base pairs in the amplified region (including the primer-targeted regions). We also used different pairs of degenerate primers (Fig. S1) for PCR isolation from total DNA and identified 16 independent *MYB* sequences. These sequences were also temporally named after the total sizes of the amplified regions (the actual sizes were 4–6 bp longer than the numbers used for the names, which were adjusted according to the targeted region of MYBdgFP1 and MYBdgRP1).

RT-PCR of MYB partial R2R3 domain from total RNA

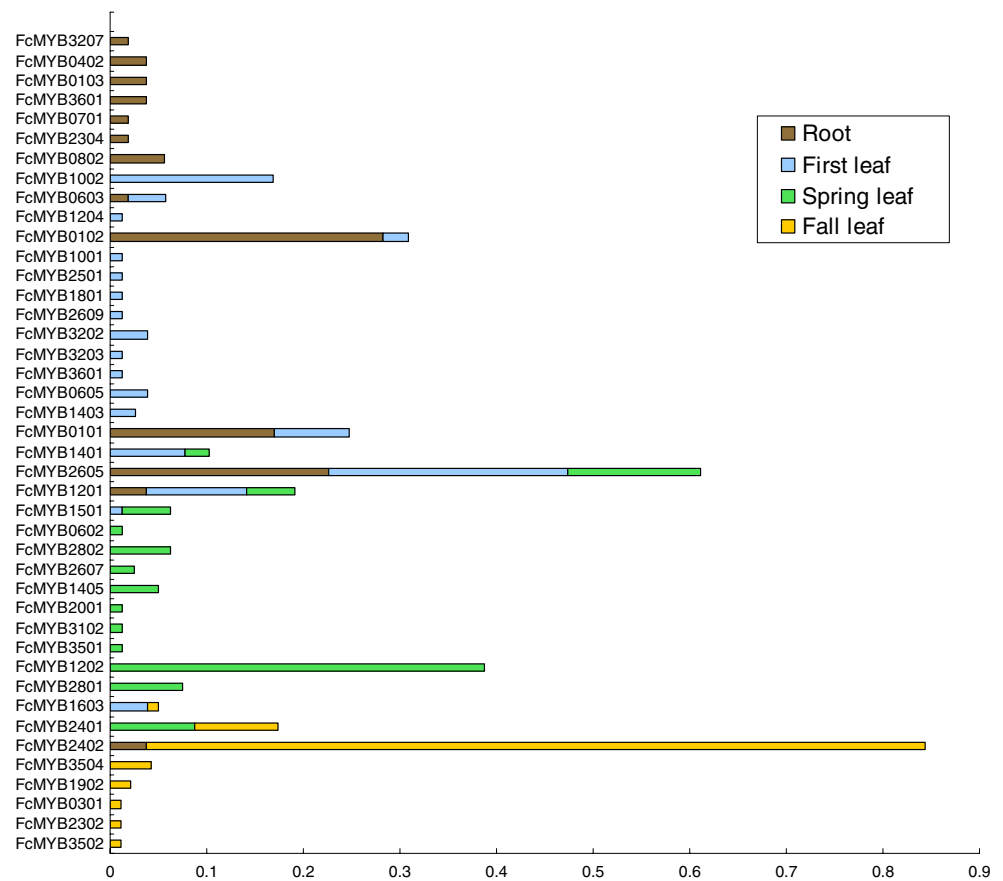
The beech *MYB* partial R2R3 domain was also amplified from total RNA prepared from the leaves sampled in fall and in spring. As expected, the RT-PCR products were detected as a single band approximately 179 bp in size; this length equaled to the total length of the primer sequences (52 bp) and the amplified region in the exons (127 bp). A variety of *MYB* cDNA sequences were identified after the

cloning and sequencing of the RT-PCR products. At this moment, all the identified members were named using capital letters derived from the names of their sample sources (F for fall and S for spring), followed by consecutive numbers, for example *FcMYBF1–8* and *FcMYBS1–12*. Because the fall members and the spring members scarcely overlapped and the only member common to both seasons was *FcMYB2401*, we speculated that many other members of the *R2R3-MYB* family might be identified from different tissues/organs under different physiological conditions. Thus, RT-PCR amplification was performed with the total RNA extracted from young spring leaves after the application of different stresses such as cold, heat, and drought as well as after gibberellic acid treatment as described in “Materials and methods”. All the identified genes were named using capital letters derived from the names of the stress conditions (C for cold, H for heat, D for drought, and GA for gibberellic acid treatment), for example, *FcMYBC1–12*, *FcMYBD1–7*, and *FcMYBH1–4*. Likewise, those identified from the GA-treated leaves (L) and petioles (P) were named *FcMYBGAL1–4* and *FcMYB-GAP1–7*, respectively. In addition, some other genes were identified from the male flower (M) and the pollen (P) and were designated *FcMYBMI–6* and *FcMYBPI–2*, respectively. We also explored new members isolated from the young seedlings (age, 1 month). Total RNA was extracted from first leaves (FL), roots (R), and stems (ST) before and after incubation of the young seedlings under cold and drought conditions for 2 h. This time, however, in order to avoid excessive complexity, we named only the novel sequences specific to the seedlings, and all the newly identified genes were designated *FcMYBFL1–34*, *FcMYBRI–24*, and *FcMYBST1*, regardless of the stress conditions. These original gene names as described above are indicated in the column of the Original name in Table S1. However, these names were replaced by more systematic names as described in “Materials and methods”.

Organ and age specificity of the identified *MYB* genes

Because we sequenced a total of 80, 93, 78, and 53 clones from the fall leaves, spring leaves, first leaves, and roots, respectively, we compared the frequency of each type of clone among the four groups. We found that the proportion of the genes identified was strikingly different depending on the organs and the age of the leaf tissues (Fig. 2). For example, *FcMYB2402* accounted for more than 80% of the total clones from fall leaves, but it was not isolated from any other source, except for the two clones from the roots. *FcMYB1202* was the most abundant in the spring leaves, but it was not isolated elsewhere. *FcMYB2605* was abundantly found from the spring leaves, first leaves, and roots, but not from the fall leaves. The member of genes

Fig. 2 Organ- and developmental stage-specific expression of the identified *MYB* genes. The frequency of clone for each member of genes from the fall leaves (*F*), spring leaves (*S*), first leaves (*FL*), and roots (*R*) is indicated as the rate of the number of clones



with lower abundance was also quite different among the four groups. Overall, 77% of all the members were isolated from only one of the four sample sources. Considered together, the above results indicate that the members of the *R2R3-MYB* family whose functions might be related to early organ development or to the process of senescence were considerably different.

Listing of all the identified *FcMYB* genes and their phylogenetic analysis

Many members with identical sequences were isolated from different sample sources, and these members were successively designated with original names. When cDNA sequences were identical to the genomic sequences in the exons, these pairs of genes were also designated according to the names of both the genomic and cDNA sequences. After integrating all the identical sequences of different origins, we documented a total of 134 independent sequences. From these sequences, we excluded 49 clones to obtain a final list of 85 genes (Table S1); these 49 clones were excluded because they differed from the genes in the final list by only one base substitution, and they were identified only once in the whole experiment. Finally, every gene in the list was systematically renamed following the

clustering results obtained by phylogenetic analysis as described below.

Phylogenetic analysis of the deduced amino acid sequences of the products of all the identified *MYB* genes along with the corresponding sequences for all 126 members of the *Arabidopsis R2R3-MYB* family revealed that a majority of the beech genes were included in the same clusters as the members from *Arabidopsis* (Kranz et al. 1998; Stracke et al. 2001) (Fig. 3). Indeed the *Arabidopsis* subgroups 1, 8, 9, 11, 13, 14, 15, 18, 19, and 22 included at least one beech gene, and these clusters were supported by bootstrap values of more than 50–90% (Fig. S2). In the case of subgroup 24, the entire cluster was not supported by a significant value, but some close relationships were found between *AtMYB53* and *FcMYB1802* and between *AtMYB93* and *FcMYB1801*. Similarly, close relationships were observed between small groups of beech and *Arabidopsis* genes in subgroups 4, 7, and 20.

We also found that a large number of beech genes were included in some expanding clusters. For instance, 23 beech genes clustered into some groups along with the *Arabidopsis* members belonging to the epidermal cell fate clade (ECFC), including subgroups 5, 6, and 15 as defined by Matus et al. (2008). In another case, nine beech genes

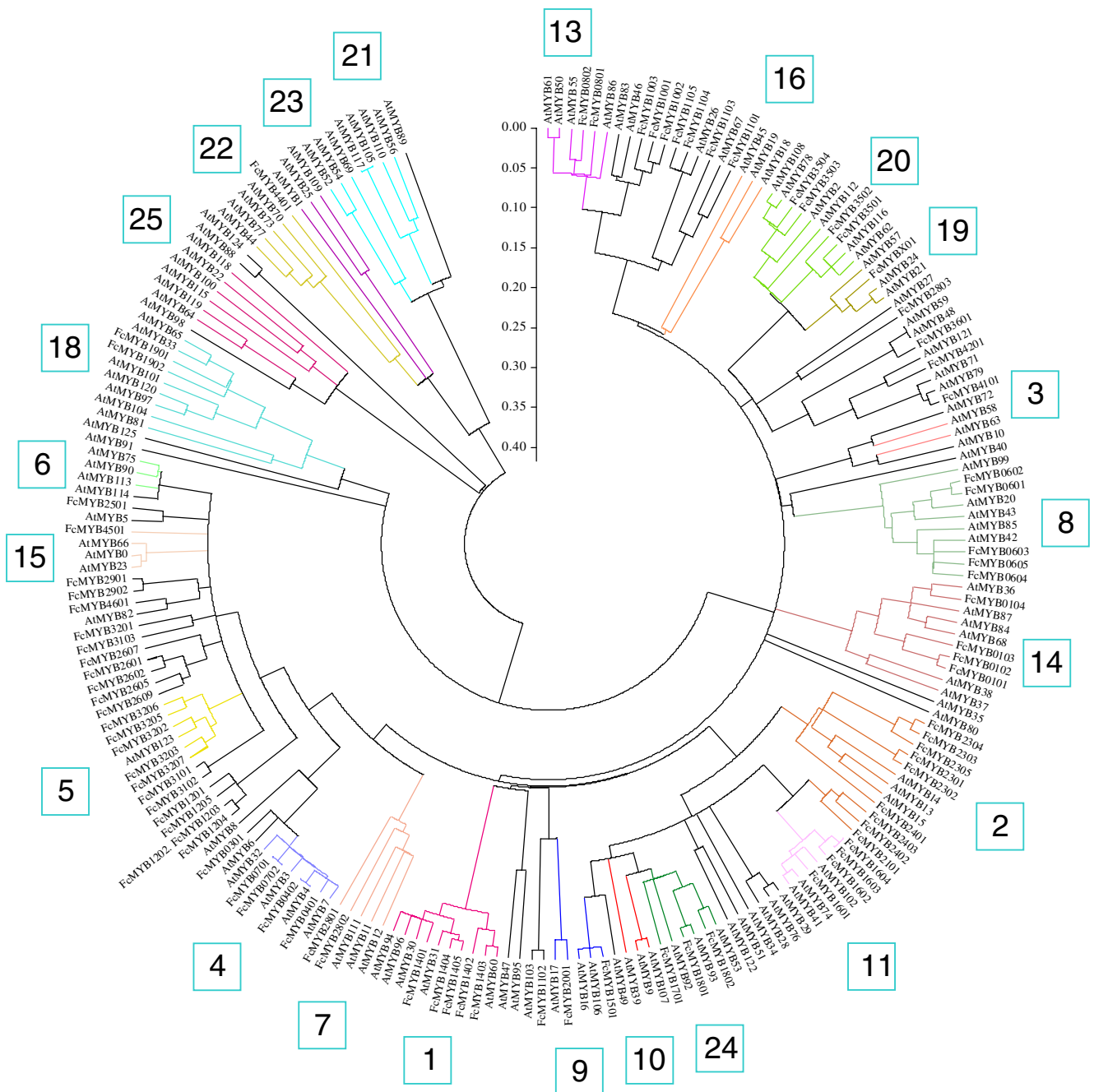


Fig. 3 Phylogenetic circular rooted tree of *Fagus*–*Arabidopsis* R2R3-MYB protein. The 42-amino-acid-long sequences of R2R3-DBD of 85 beech members were compared with the corresponding sequences for 126 members of the *Arabidopsis* R2R3-MYB family to construct a tree

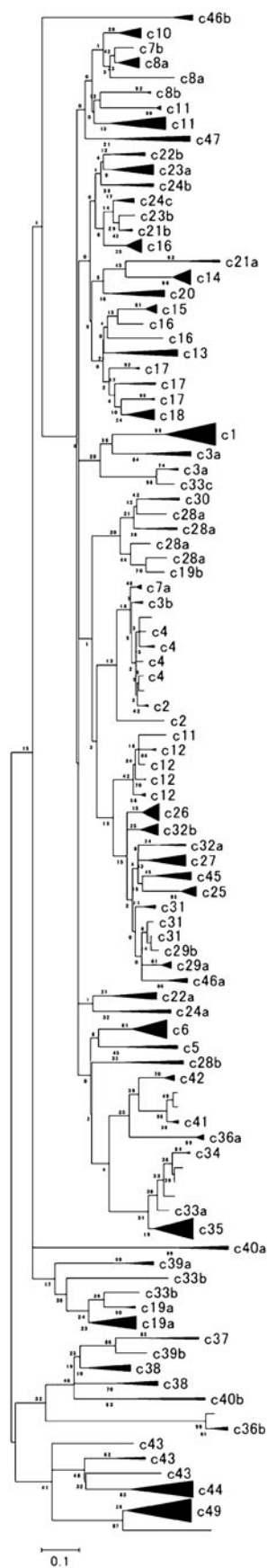
with Mega4 software using the neighbor-joining (NJ) method and 5,000 bootstraps. Subgroup designations of beech members were based on the tree by following Kranz et al. (1998) and Stracke et al. (2001). The beech genes with synonyms were represented by the first gene name

formed a cluster that showed a close relationship with the *Arabidopsis* members of subgroup 2.

The presence of expanding clusters implied that these genes might have been amplified in the beech genome. To investigate how these beech members in the expanding clusters were related to the genes from the *Populus* R2R3-MYB family, another phylogeny was constructed using the corresponding 42-amino-acid-long sequences of the prod-

ucts encoded by all the *Populus* members in addition to those of the products encoded by the beech and *Arabidopsis* members (Fig. 4). In terms of cluster formation, the tree was very similar to the published phylogeny constructed with all the *Populus*, *Arabidopsis*, and *Vitis* R2R3-MYB family members (Wilkins et al. 2009). Some of the above clades were subdivided into different clusters and were distinguished with an additional alphabet, such as 24a, 24b,

Fig. 4 Phylogenetic relationships and subgroup designations in 85, 192, and 126 members of the beech, *Populus*, and *Arabidopsis* R2R3-MYB families, respectively, based on 42-amino-acid-long sequences of R2R3-DBD. The tree was constructed as described in the caption to Fig. 3. All these families' members were categorized into 49 clades (triangles) by following Wilkins et al. The uncompressed tree with full taxa names is available as Supplementary Fig. S3. Some of the clades were subdivided into different clusters and were distinguished with an alphabet, such as 24a, 24b, and 24c

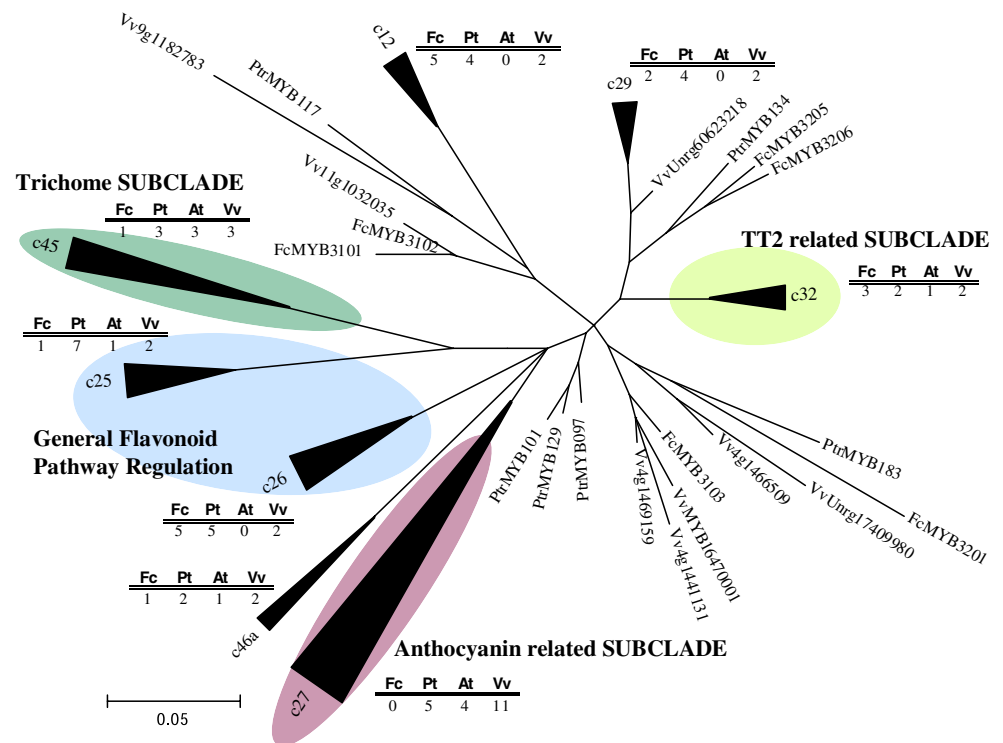


and 24c. At least one of the beech members was found in 32 of the 49 clades defined in the reported phylogeny (Table S2). In clades such as C1, 4, 6, 8, 10, 11, 14, 15, 16, 17, 19, 20, 21, 28, 35, 36, 41, 42, and 44, beech genes clustered both with *Arabidopsis* and *Populus* genes, and in many cases, the beech genes showed a closer relation to the *Populus* genes than to the *Arabidopsis* genes (Fig. S3). Extreme results were found in the cases of clades C12, 23, 24, 26, 29, 31, and 32; these clades included a total of 25 beech members and *AtMYB123* (TT2) in C32 was the only *Arabidopsis* included gene (Table S2). Thus, the trends of gene expansion in these clades reported for the *Vitis* and *Populus* families seem to be conserved in the beech family.

Phylogenetic analysis of the beech genes in the ECFC

Since a majority of the beech expanding clusters related to the *Arabidopsis* ECFC were found in clades C12, 26, 29, 31, and 32 of the *Populus* and *Vitis* family, wherein expansion was also reported, we constructed another phylogenetic tree with all the *Vitis* genes belonging to the ECFC, in addition to the corresponding members in the *Arabidopsis*, *Populus*, and beech families. As shown in Fig. 5 (and Fig. S4), the whole tree was constructed with members of clades C12, 25, 26, 27, 29, 31, 32, 45, and 46a. Four subclades, designated according to the characteristic function of their representative genes, were similarly constructed as reported before (Matus et al. 2008). The subclade “general flavonoid pathway regulation” consisted of members of C25 and C26. Beech genes were abundant in C26, and all the five members in this clade formed a cluster with *Vitis MYBPA1* and GSVIVT00006679001. *VvMYBPA1* has been reported to regulate proanthocyanidin (PA) synthesis during the development of grape berries. C25 included *VvMYB5a*, *VvMYB5b*, and *AtMYB5*. *VvMYB5a* and *VvMYB5b* have also been described as regulators of the genes involved in flavonoid biosynthesis. C32 contained many beech genes, and three of six genes were included in a branch of a TT2-related subclade. In contrast to these subclades, beech genes were scarcely found in the other two subclades. Only one gene was found in the trichome subclade, and none in the anthocyanin-related subclade. Because it was possible that the degenerate primers used in the general PCR experiments were not suitable for the isolation of genes in the anthocyanin-related subclade, we designed another pair of primers on the basis of the consensus sequences of *Populus* and *Arabidopsis MYB* genes in the anthocyanin-related subclade (AnFP1 and AnRP1, depicted in Fig. S1). Using these specific primers, we identified four types of additional *MYB* sequences from total DNA and named them *FcMYB1103*, *FcMYB1104*, *FcMYB1105*, and *FcMYB0702*. However, none of these sequences appeared

Fig. 5 Phylogenetic analysis of R2R3-MYB in the epidermal cell fate clade. There were 23, 32, 38, and 10 members of MYB amino acid sequences of the products of all members from the beech, *Arabidopsis*, *Populus*, and *Vitis* families, respectively, that were compared based on 42-amino-acid-long sequences of R2R3-DBD to construct an unrooted tree as described in the caption to Fig. 3. Subgroup designations for the compressed triangles refer to the tree in Fig. 4. The numbers of genes from the four species are indicated beside the triangles. The uncompressed tree with full taxa names is available as Supplementary Fig. S4



in the anthocyanin-related subclade but were included in clades C7 and C11.

It should be kept in mind that these phylogenetic analyses are constructed based on a small portion of the genes and in some cases they are classified with a small number of polymorphism. Thus, both the expansion of general flavonoid pathway MYBs and the absence of anthocyanin MYBs have to be considered very carefully: the flavonoid pathway MYBs might include some extra members that should have been placed in a different clade and some potential anthocyanin members might be being misplaced in a different clade nearby. In order to clarify this possibility, we have to determine the three-terminal portion of all these genes followed by functional analysis.

Complete structures of genes in expanding clusters and putative orthologs

As indicated in the phylogenetic analysis, many of the beech genes showed orthologous relationships with members of the *Populus* and *Arabidopsis* families. One such member, designated *FcMYB1901*, was grouped in the same cluster as the *GAMYB* genes, e.g., *AtMYB33*, *65*, and *101*. We isolated a lambda genomic clone containing the complete coding region of *FcMYB1901* and determined its nucleotide sequence. It was found that *FcMYB1901* contained every motif characteristic of the group of genes orthologous to *HvGAMYB* (Gocal et al. 2001; Achard et al. 2004; Millar and Gubler 2005), for example, the coding sequences for boxes 1, 2, and 3, and the putative

recognition sequence of miR159 (Table 1). The expression profile of *FcMYB1901* also showed some characteristic features of *GAMYB* in that this gene was strongly expressed in the anther and pollen but was barely detectable in the leaves and female flowers (Fig. S5). When twigs with winter buds were incubated at 25°C, the emergence of young leaves and elongation of the petioles were evidently promoted by the addition of 0.1 mM gibberellin during incubation. Further, *FcMYB1901* was upregulated by the gibberellin treatment (Fig. S5). All these results suggest that *FcMYB1901* is a candidate ortholog of *GAMYB* in beech.

Another example of an orthologous relationship was suggested between *FcMYB2402* (originally named as 828/*F1*) and the genes in subgroup 2, including *AtMYB15*. *FcMYB2402* was identified as one of the most abundantly expressed MYB genes in the fall leaves and was not found among the spring leaves. By using the *FcMYBF1* cDNA fragment as a probe, we isolated three positive clones from the lambda genomic library and determined the nucleotide sequences of their coding regions. As expected, these three clones contained exon and intron sequences that were identical to those of *FcMYB828*. Thus, we concluded that *FcMYBF1* was a transcribed sequence of *FcMYB828*. The deduced amino acid sequence of the product of *FcMY2402* was aligned with the sequences of the products of all the members of clades C22, 23, and C24 from the *Populus*, *Arabidopsis*, and *Vitis* families (Fig. 6). The alignments revealed that R2R3-DBD was very similar to all the members in these clades. In contrast, the amino acid motifs characteristic for subgroup 2 defined as [I/M]DExFWS[D/

Table 1 Structural motifs characteristic of GAMYB

Gene	BOX1	BOX2	miR159 target sequence	BOX3
<i>HvGAMYB</i>	QRAGLP-IYPAS	NSGLLEELLHEAQ	ATGGAGCTCCCTTCACTCCAAG	WSNLP
<i>AtMYB33</i>	QRAGLP-LYPPE	DTGLLDMLLEAK	CTGGAGCTCCCTTCAITCCAAT	WSNMP
<i>AtMYB65</i>	QRAGLP-LYPPE	DTGLLDMLLHEAK	CTGGAGCTCCCTTCAITCCAAT	WTNMP
<i>AtMYB101</i>	QRAGLP-LYPHE	NSGLLDALLEESQ	CTAGAGCTTCCTTCAAACCAA	WSNMP
<i>FcMYB1901</i>	QRQGLPPLYPHD	GNGLLEDLFDEAH	CACGAGCTCCCTTCAAGCCAAT	WDNMP
Consensus	QRaGLP-LYPxe	xxGLLxxLlxEax	X E L P S X Q	WxNmP

E] and [D/N]xxM[D/E]FW[Y/F/H][D/N][V/L/I]F were shared only by a group of genes, including three members each of the *Populus* and *Arabidopsis* and four members of the *Vitis* family in C22b, C23b, and C24c. Very similar motifs were also found in *FcMYB2402*, though the first motif was partially diverged. These motifs were scarcely conserved in the other members of the *Populus* and *Vitis* families in C23a, C24a, and C24b. This indicated that the reported expansion of the *Populus* and *Vitis* families in C23 and C24 was mainly ascribed to the members without the C-terminal motifs in subgroup 2. In the case of the beech family, five genes, including *FcMYB2301* and *FcMYB2302*, showed close relationships with four members of the *Populus* family in C23a; *FcMYB2401* formed a cluster with *PtrMYB049* and *PtrMYB063* in C24b, and *FcMYB2101*, which also belonged to subgroup 2 with *Arabidopsis* genes, formed another cluster with *PtrMYB058* in C21b (Fig. S3). None of these *Populus* members contained C-terminal motifs, and only *FcMYB2402* and *FcMYB2403* clustered with the four *Populus* members containing C-terminal motifs in C24c (Fig. 6; Fig. S3). Thus, among the nine beech genes in subgroup 2-related expanding cluster, only the *FcMYB2402* and *FcMYB2403* seemed to be putative orthologs of the genes in subgroup 2 of *Arabidopsis*. In addition to this structural similarity, *FcMYB2402* also showed signs of functional similarity: the expression of this gene was rapidly upregulated upon cold treatment of the emerging young spring leaves (Fig. S6).

In order to analyze some members of the ECFC in greater detail, we isolated genomic clones containing *FcMYB3202*, *FcMYB3201*, and *FcMYB3103* and determined their complete sequences. In Fig. 5, *FcMYB3202* was included in TT2-related subclade, and *FcMYB3201* and *FcMYB3103* were found outside of any functionally defined subclades. The structural analysis of these genes revealed that the amino acid sequence of the product encoded by *FcMYB3202* exhibited a motif that was very similar to that of TT2 (VIRTKAxRC[T/S]KxL); these motifs were located 43 or 25 amino acid residues downstream of the R2R3 domain, respectively (Fig. S7). TT2 functions as a transcription factor for the genes involved in the pathway of proanthocyanin biosynthesis,

such as *dihydroflavonol-4-reductase* and *BANYULS*, and is required for seed coat pigmentation (Nesi et al. 2001). *FcMYB3202* was also closely related to cotton genes such as *GhMYB10*, *GhMYB36*, and *GhMYB38* (data not shown), which were found to be expressed in the ovules (Cedroni et al. 2003). A similar motif was also found in *ZmMYBC1*, which has been reported to control the anthocyanin structural genes in maize (Cone et al. 1986), suggesting that this motif might be required for the activation of the pathways of proanthocyanin and anthocyanin biosynthesis by TT2 and *ZmMYBC1*, respectively. On the basis of all these related sequences, the C-terminal core motif was estimated as V x₃KAxRC[T/S]. Interestingly, although *FcMYB3201* was found to be relatively distinct from TT2 in the phylogenetic analysis, it also shared a similar motif with TT2. However, this motif was not found in *FcMYB3103*, though the sequence of *FcMYB3103* seemed to be closest to that of *FcMYB3201* (Fig. 5).

FcMYB3103 was found to be tandemly repeated in two copies, separated by a 1,501-bp-long intergenic region, between the two open reading frames (ORFs) of *FcMYB3103-1* and *FcMYB3103-2*. When the two duplicated sequences were compared, it was found that the mutation rate was not much higher in the non-translated regions than in the ORF regions. In fact, the ORF variable region (VR) was more prone to mutation than the 5' non-translated region. It was also found that nonsynonymous (N) mutations were three times more frequent than synonymous (S) mutations in VR that was compatible to the potential N/S ratio, whereas N and S appeared equally in the ORF DNA binding domain (DBD) that was much lower than the potential N/S ratio of 3.31 (Table 2). Thus, the results indicated that the mutations in VR had accumulated without purifying selection, while some selective pressure had been applied to DBD.

Discussion

R2R3-MYB regulatory genes are a large family of more than 100 members and are found in many higher plants such as *A. thaliana*, *P. trichocarpa*, and *V. vinifera*. It has been

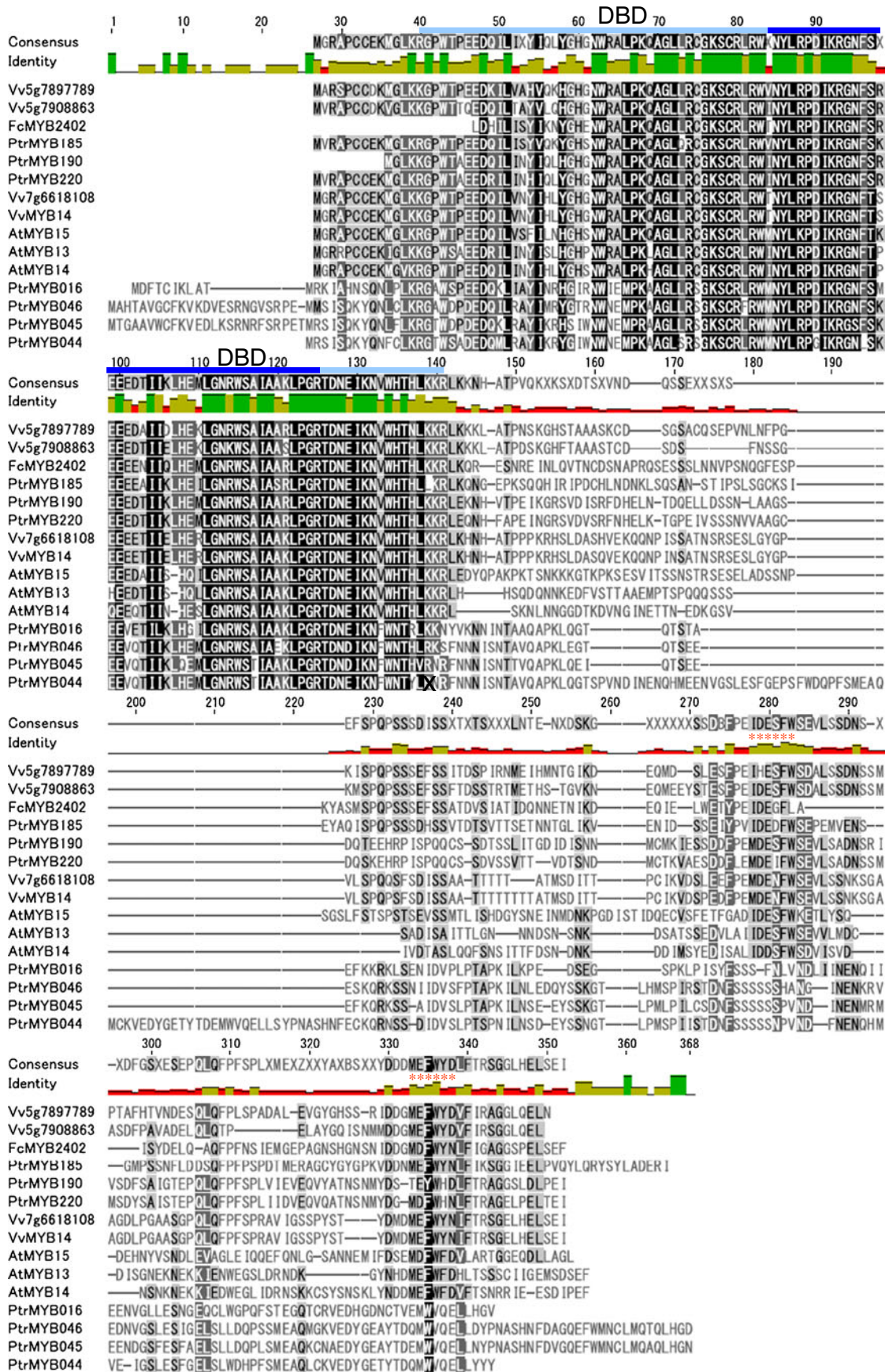


Fig. 6 Sequence alignment of the MYB proteins encoded by genes in the expanding clade related to subgroup 2. Alignment of the full-length amino acid sequences of the products of *AtMYB15* and its related members from the beech, *Arabidopsis*, *Populus*, and *Vitis* families respectively, and the PDF images of the alignments were generated using Clustal W program of Geneious v4.5.3. Residues with various shades of background indicate the similarity levels: black, 100%; dark gray, 80% to 100%; gray, 60% to 80%; and white, less than 60% identical among all the proteins. The whole region of R2R3-DBD is indicated with a bar in light blue above consensus residues with the 42-a.a. region used for the phylogenetic analysis colored in dark blue. Red asterisks indicate C-terminal motives

demonstrated that these genes are involved in a wide range of biological processes. Some of these members function in response to various types of environmental stresses such as cold, high temperature, drought, and nutritional limitation. In order to identify the genes in beech that function in response to environmental changes, we first attempted to collect as many genes belonging to the *R2R3-MYB* family as possible. Pairs of degenerate primers were used to amplify the *MYB* partial R2R3 domain coding for a 42-amino-acid-long product from both total DNA and total RNA. Thus far, we have identified the partial sequences of at least 85 independent genes. Although we have probably not yet isolated all the members of the beech *R2R3-MYB* family, we seem to have identified a substantial number of candidate genes relevant to environmental acclimation. Phylogenetic analysis of the identified sequences and further isolation of the complete coding region of some representative genes revealed some interesting features of the beech *R2R3-MYB* family.

Features of the beech *R2R3-MYB* family common to other plants

In the phylogeny constructed with all the identified beech genes containing the MYB partial R2R3 domain and all the *Arabidopsis* family members containing a corresponding domain, many of the beech genes formed clusters with at least one member of the *Arabidopsis* subgroups. Some of the orthologous relationships were more evident during the comparison of the complete structures of the above genes.

For example, *FcMYB1901* exhibited all the characteristic motifs of *GAMYB* orthologs. The structure of *FcMYB3202* was also very similar to that of *AtMYB123*, and that of *FcMYB2402* was similar to the structures of three genes in *Arabidopsis* subgroup 2. These results will be helpful to estimate similar types of orthologous relationships and narrow down relevant candidate genes. The above experimental data reveal some interesting relationships, as listed in Table 3. For example, *FcMYB1603* shows an orthologous relationship with the members of *Arabidopsis* subgroup 11, including *AtMYB41* and *AtMYB102*. While it has been reported that the latter two genes are involved in osmotic stress responses in *Arabidopsis*, we observed that the frequency of *FcMYB1603* clones identified from first leaves and roots of beech young seedlings increased after drought treatment (data not shown). Thus, the expression of the gene might be induced by osmotic stress resulting from drought conditions. Clones of *FcMYB1002* were abundantly isolated from the first leaves under normal conditions but were considerably reduced after drought and cold treatment, respectively. This gene is a candidate ortholog of *AtMYB46*, which regulates lignin biosynthesis in the secondary wall-associated *nac domain protein1*-dependent transcriptional network (Zhong et al. 2007). It is intriguing to speculate that *FcMYB1002* is induced to activate secondary wall biosynthesis under normal growing conditions and is repressed under stress conditions for the immediate cessation of growth. In either case, the function of these beech genes could be further demonstrated using complementation experiments with *Arabidopsis* mutants.

Expansion of the beech *R2R3-MYB* family in the ECFC

A large number of beech genes form some groups of expanding clades, which are more or less related to the *Arabidopsis* members of subgroups 5, 6, and 15, which are included in the ECFC. In grape and poplar, the number of genes belonging to these clades is increased, and in beech a comparable number of genes is found in the general flavonoid pathway regulation and TT2-related subclades.

Table 2 Mutation rates of the recently duplicated copies of *FcMYB3103-1* and *3103-2*

	5' non-translated region	5' DNA binding domain	3' variable region 3103-1/3103-2	Introns 1 and 2	3' non-translated region
No. of base pairs	640	345	393/396	229	238
No. of SNP + in/del	13	6	12	8	12
Mutation rate	2.03	1.92	3.03/3.05	3.49	5.04
Nonsynonymous (potential sites)	–	3 (265)	9 (303/304.75)	–	–
Synonymous (potential sites ^a)	–	3 (80)	3 (90/91.25)	–	–
N/S (potential sites ^a)	–	1.0 (3.31)	3.0 (3.37/3.34)	–	–

^a Calculated according to Zhang et al. (1998), assuming that the rates of transitional and transversional changes are the same

Table 3 Putative orthologs of the beech genes and their estimated functions

Bootstrap value	Beech genes	Putative ortholog	Subgroup	Clade	Estimated functions	References
65*2	<i>FcMYB1403</i>	<i>AtMYB60</i>	1	14	Stomatal movement	Cominelli et al. 2005
38*1	<i>FcMYB0402</i> <i>FcMYB0401</i>	<i>AtMYB4</i>	4	4	Repressor of C4H, normal pollen development	Preston et al. 2004
38*1	<i>FcMYB3202</i>	<i>AtMYB123</i>	5	32b	Seed coat pigmentation	Nesi et al. 2001
37*1	<i>FcMYB2802</i>	<i>AtMYB11</i> <i>AtMYB12</i> <i>AtMYB111</i>	7	28a	Tissue-specific flavonoid accumulation, modulate meristematic activity	Stracke et al. 2007 Petroni et al. 2008
89*1	<i>FcMYB1501</i>	<i>AtMYB16</i> <i>AtMYB106</i>	9	15	Cell and petal morphogenesis, leaf hair development	Baumann et al. 2007
92*1	<i>FcMYB2001</i>	<i>AtMYB17</i>	9	20	Possible regulation by LEAFY and AGL15 for early inflorescence development	Zhang et al. 2009
76*1	<i>FcMYB1601</i> <i>FcMYB1602</i> <i>FcMYB1603</i> <i>FcMYB1604</i>	<i>AtMYB41</i> <i>AtMYB74</i> <i>AtMYB102</i>	11	16a	Cell expansion, cuticle development, osmotic and wound response	Denekamp and Smeekens 2003 Cominelli et al. 2008 Lippold et al. 2009
69*1	<i>FcMYB0801</i> <i>FcMYB0802</i>	<i>AtMYB50</i> <i>AtMYB55</i> <i>AtMYB61</i> <i>AtMYB86</i>	13	8a, 7b	Dark photomorphogenesis and control of stomatal aperture	Newman et al. 2004 Liang et al. 2005
97*1	<i>FcMYB0101</i> <i>FcMYB0102</i> <i>FcMYB0103</i> <i>FcMYB0104</i>	<i>AtMYB36</i> <i>AtMYB37</i> <i>AtMYB38</i> <i>AtMYB68</i> <i>AtMYB84</i> <i>AtMYB87</i>	14	1	Regulator of axillary meristem	Schmitz et al. 2002 Müller et al. 2006 Kellar et al. 2006
64*1	<i>FcMYB4501</i>	<i>AtMYB0</i> <i>AtMYB23</i> <i>AtMYB66</i>	15	45	Epidermal cell patterning	Oppenheimer et al. 1991
54*1	<i>FcMYB1901</i> <i>FcMYB1902</i>	<i>AtMYB33</i> <i>AtMYB65</i> <i>AtMYB97</i> <i>AtMYB101</i> <i>AtMYB120</i>	18	19a	To mediate GA signaling for growth and flowering	Gubler et al. 1995 Gocal et al. 2001
60*1	<i>FcMYBX01</i>	<i>AtMYB21</i> <i>AtMYB24</i>	19		Flower-specific transcription factor	Shin et al. 2002
60*1	<i>FcMYB3503</i> <i>FcMYB3504</i>	<i>AtMYB2</i> <i>AtMYB78</i> <i>AtMYB108</i>	20	35	Induction of ADH1 by low oxygen in ABA signaling, drought and salt tolerance	Hoeren et al. 1998 Abe et al. 2003 Yoo et al. 2005
82*1	<i>FcMYB4401</i>	<i>AtMYB44</i> <i>AtMYB70</i> <i>AtMYB73</i> <i>AtMYB77</i>	22	44	Auxin signal transduction, stomatal closure	Shin et al. 2007 Jung et al. 2008
99*2	<i>FcMYB1102</i>	<i>AtMYB103</i>		11	Regulate tapetum development, callose dissolution, and exine formation	Zhang et al. 2007
33*1	<i>FcMYB1001</i> <i>FcMYB1002</i> <i>FcMYB1003</i>	<i>AtMYB46</i> <i>AtMYB83</i>		10	Secondary wall biosynthesis	Zhong et al. 2007

*1, *2 the values indicated in Supplemental Fig. S2 and S3, respectively

In the case of grape vines, because PA and condensed tannins are important to ensure the astringency and bitterness of wine, the processes for the cultivation of grape vines may have facilitated the diversity of the genes in these subclades. On the other hand, expansion of the beech gene family in the ECFC might have contributed to the evolution of a natural tree species by the addition of new phases to the physiology of flavonoid metabolism.

Flavonoids function as biotic and abiotic protectants, such as phytoalexins and UV-B protectants. In leguminous plants, specific types of flavonoids serve as a cue to the formation of nitrogen-fixing root nodules. Moreover, flavonols such as kaempferol and quercetin function as competitive regulators of auxin transport (Buer et al. 2010). Though these physiologically active natures of flavonoids have not been well studied in beech, a broad possibility could be considered in the physiological roles for the beech expanding flavonoid MYB genes and it may be advantageous to have additional copies of genes with redundant function to gain a more complicated form of life and versatile means of survival.

The nucleotide sequences of *FcMYB3202*, *FcMYB3103*, and *FcMYB3201* are very similar to the sequence of *AtMYB123* (*TT2*) in the R2R3-DBD, while *FcMYB3202* and *FcMYB3201* share a conserved motif with *TT2* in the C-terminal region; this motif is not found in *FcMYB3103*. These genes are also similar to some cotton genes, namely, *GhMYB10*, *GhMYB36*, and *GhMYB38* (accession numbers AF336282, AF336284, and AF336285, respectively), which have been found to be expressed in the ovules. These results suggest that some members of the expanding cluster have maintained their basic functions and, in addition, acquired some new functions and thereby contribute to the alternative method of survival in nature. A similar evolutionary process has been reported for the cotton gene, *GaMYB2*, which performs the function of trichome formation in *Arabidopsis gll* mutant leaves and also induces seed trichome production, suggesting that this gene may have gained a new function as a regulator of cotton fiber development while maintaining its basic function as a regulator of leaf trichome formation (Wang et al. 2004).

The duplicated genes of *FcMYB3103-1* and *FcMYB3103-2* may provide an interesting example of an evolutionary intermediate; their sequence similarity is very high, suggesting that the duplication must have been a relatively recent event. Indeed the extensive sequence similarity was observed even in the non-coding regions, yet these twin genes have already started to diverge substantially. Importantly, because the N/S ratio in their DBDs is lower than the potential N/S ratio, purifying selection may have occurred against some N mutations in this region, suggesting that the activation of both these genes in the beech genome is advantageous. It is also possible that these genes developed differential roles in beech physiology. Further analysis of the functions of the

newly amplified regulatory genes as well as the beech-specific cluster genes may clarify the roles of these genes in beech evolution.

The *Arabidopsis* members of the ECFC, such as *GLI* (*AtMYB0*), *WER* (*AtMYB66*), *PAP1* (*AtMYB75*), *PAP2* (*AtMYB90*), *TT2*, and *AtMYB5*, are involved in the transcriptional control of epidermal cell differentiation, including trichome morphogenesis, root hair formation, anthocyanin production, and seed coat pigmentation and development. Because all these members function in combination with bHLH proteins, the amino acid sequence motif required for their interaction ($[D/E]Lx_2[R/K]x_3Lx_6Lx_3R$) is very well conserved in the second repeat of the R2R3-DBD (Tominaga et al. 2007). Interestingly, the beech members included here in the ECFC are distinguished by the fact that they all possess an interaction motif, while this motif is not present in any other members of the entire beech family that have been thus far identified, except for *FcMYB0402* and *FcMYB0401* in C4. Thus, all the beech members in the ECFC have a conserved structural basis for interaction with a bHLH protein; this interaction is necessary to form a functional protein complex for epidermal cell differentiation.

Diversion of the expanding members related to subgroup 2

In the case of the expanding clade of beech closely related to *Arabidopsis* subgroup 2, the members were clearly subdivided into the two classes. The first class included *FcMYB2402* and possibly *FcMYB2403* and was distinguished by the presence of C-terminal motifs common to the genes of subgroup 2. The second class of beech genes included seven members and was closer to the *Populus* members without the C-terminal motifs. In light of the above information, fluctuation of the number of genes in the *Arabidopsis*, *Populus*, and *Vitis* families in clades C22, C23, and C24 may be viewed differently. This is because the number of genes in the first class was more or less consistent among these plant species. In contrast, no *Arabidopsis* genes were included in the second class, and the number of genes belonging to this class varied among the beech, poplar, and grape families. These genes may be another example of evolutionary intermediates, similar to the twin genes of *FcMYB3103*. If the physiological role of these genes in these species is not yet fixed, various alleles of these genes might exist, and the allelic frequency may fluctuate depending on geographic area and climatic conditions. It would be intriguing to assume that these genes are still evolving or that they act as potential genetic resources for the changing environment owing to their highly heterogeneous state in the beech population. Further investigation into the genetic diversity and functional analysis of these genes are required to confirm these speculations.

Funding This study was supported by a Grant-in-Aid for Scientific Research (no. 15380219) from Japan Society for Promotion of Sciences (JSPS).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Arabidopsis AtMYC2(bHLH) and AtMYB2(MYB) function as transcriptional Activators in abscisic acid signaling. *Plant Cell* 15:63–78
- Achard P, Herr A, Baulcombe DC, Harberd NP (2004) Modulation of floral development by a gibberellin-regulated microRNA. *Development* 131:3357–3365
- Asuka Y, Tomaru N, Nisimura N, Tsumura Y, Yamamoto S (2004) Heterogeneous genetic structure in a *Fagus crenata* population in an old-growth beech forest revealed by microsatellite markers. *Mol Ecol* 13:1241–1250
- Asuka Y, Tomaru N, Munehara Y, Tani N, Tsumura Y, Yamamoto S (2005) Half-sib family structure of *Fagus crenata* saplings in an old-growth beech–dwarf bamboo forest. *Mol Ecol* 14:2565–2575
- Baumann K, Perez-Rodriguez M, Bradley D, Venail J, Bailey P, Jin H, Koes R, Roberts K, Martin C (2007) Control of cell and petal morphogenesis by R2R3 MYB transcription factors. *Development* 134:1691–1701
- Bender J, Fink GR (1998) A Myb homologue, ATR1, activates tryptophan gene expression in *Arabidopsis*. *Proc Natl Acad Sci USA* 95:5655–5660
- Buer CS, Imin N, Djordjevic MA (2010) Flavonoids: new roles for old molecules. *J Integr Plant Biol* 52:98–111
- Cedroni ML, Cronn RC, Adams KL, Wilkins TA, Wendel JF (2003) Evolution and expression of MYB genes in diploid and polyploid cotton. *Plant Mol Biol* 51:313–325
- Chang S, Puryear J, Cairney J (1993) Simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11:113–116
- Cominelli E, Galbiati M, Vavasseur A, Conti L, Sala T, Vuylsteke M, Leonhardt N, Dellaporta SL, Tonelli C (2005) A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. *Curr Biol* 15:1195–1200
- Cominelli E, Sala T, Calvi D, Gusmaroli G, Tonelli C (2008) Overexpression of the Arabidopsis AtMYB41 gene alters cell expansion and leaf surface permeability. *Plant J* 53:53–64
- Comps B, Gömöry D, Letouzey J, Thiébaud B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* 157:389–397
- Cone KC, Burr FA, Burr B (1986) Molecular analysis of the maize anthocyanin regulatory locus Cl. *Proc Natl Acad Sci USA* 83:9631–9635
- Denekamp M, Smeekens SC (2003) Integration of wounding and osmotic stress signals determines the expression of the *ATMYB102* transcription factor gene. *Plant Physiol* 132:1415–1423
- Gocal GFW, Sheldon CC, Gubler F, Moritz T, Bagnall DJ, MacMillan CP, Li SF, Parish RW, Dennis ES, Weigel D, King RW (2001) *GAMYB-like* genes, flowering, and gibberellin signaling in *Arabidopsis*. *Plant Physiol* 127:1682–1693
- Gubler F, Kalla R, Roberts JK, Jacobsen JV (1995) Gibberellin-regulated expression of a myb gene in barley aleurone cells: evidence for Myb transactivation of a high-pI alpha-amylase gene promoter. *Plant Cell* 7:1879–1891
- Higginson T, Feng S, Parish RW (2003) AtMYB103 regulates tapetum and trichome development in *Arabidopsis thaliana*. *Plant J* 35:177–192
- Hoeren FU, Dolferus R, Wu Y, Peacock WJ, Dennis ES (1998) Evidence for a role for AtMYB2 in the induction of the *Arabidopsis* alcohol dehydrogenase gene (*ADH1*) by low oxygen. *Genetics* 149:479–490
- Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *EMBO J* 19:6150–6161
- Jung C, Seo JS, Han SW, Koo YJ, Kim CH, Song SI, Nahm BH, Choi YD, Cheong JJ (2008) Overexpression of AtMYB44 Enhances Stomatal Closure to Confer Abiotic Stress Tolerance in Transgenic Arabidopsis. *Plant Physiol* 146:623–635
- Kellar T, Abbott J, Moritz T, Doemer P (2006) *Arabidopsis REGULATORY OF AXILLARY MERISTEMS1* controls a leaf axil stem cell niche and modulates vegetative development. *Plant Cell* 18:598–611
- Kirik V, Kolle K, Wohlfarth T, Misera S, Baumlein H (1998) Ectopic expression of a novel MYB gene modifies the architecture of the *Arabidopsis* inflorescence. *Plant J* 13:729–742
- Kranz HD, Denekamp M, Greco R, Jin H, Leyva A, Meissner RC, Petroni K, Urzainqui A, Bevan M, Martin C, Smeekens S, Tonelli C, Paz-Ares J, Weisshaar B (1998) Towards functional characterization of the members of the *R2R3-MYB* gene family from *Arabidopsis thaliana*. *Plant J* 16:263–276
- Liang YK, Dubos C, Dodd IC, Holroyd GH, Hetherington AM (2005) *AtMYB61*, an R2R3-MYB transcription factor controlling stomatal aperture in *Arabidopsis thaliana*. *Curr Biol* 15:1201–1206
- Lippold F, Sanchez DH, Musialak M, Schlereth A, Scheible WR, Hincha DK, Udvardi MK (2009) *AtMyb41* regulates transcriptional and metabolic responses to osmotic stress in *Arabidopsis*. *Plant Physiol* 149:1761–1772
- Matus JT, Aquea F, Arce-Johnson P (2008) Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across *Vitis* and *Arabidopsis* genomes. *BMC Plant Biol* 22:8–83. doi:10.1186/1471-2229-8-83
- Millar AA, Gubler F (2005) The *Arabidopsis* GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell* 17:705–721
- Müller D, Schmitz G, Theres K (2006) *Blind* homologous *R2R3 myb* genes control the pattern of lateral meristem initiation in *Arabidopsis*. *Plant Cell* 18:586–597
- Nesi N, Jond C, Debeaujon I, Caboche M, 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. *Plant Cell* 13:2099–2114
- Newman LJ, Perazza DE, Juda L, Campbell MM (2004) Involvement of the *R2R3-MYB*, *AtMYB61*, in the ectopic lignification and dark-photomorphogenic components of the *det3* mutant phenotype. *Plant J* 37:239–250
- Oppenheimer DG, Herman PL, Sivakumaran S, Esch J, Marks MD (1991) A *myb* gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. *Cell* 67:483–493
- Park JS, Kim JB, Cho KJ, Cheon CI, Sung MK, Choung MG, Roh KH (2008) *Arabidopsis* R2R3-MYB transcription factor AtMYB60 functions as a transcriptional repressor of anthocyanin biosynthesis in lettuce (*Lactuca sativa*). *Plant Cell Rep* 27:985–994

- Paz-Ares J, Ghosal D, Wienand U, Peterson PA, 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. *EMBO J* 6:3553–3558
- Petroni K, Falasca G, Calvenzani V, Allegra D, Stolfi C, Fabrizi L, Altamura MM, Tonelli C (2008) The *AtMYB11* gene from *Arabidopsis* is expressed in meristematic cells and modulates growth in planta and organogenesis in vitro. *J Exp Bot* 59:1201–1213
- Preston J, Wheeler J, Heazlewood J, Li SF, Parish RW (2004) *AtMYB32* is required for normal pollen development in *Arabidopsis thaliana*. *Plant J* 40:979–995
- Raffael S, Vaillau F, Léger A, Joubès J, Miersch O, Huard C, Blée E, Mongrand S, Domergue F, Roby D (2008) A MYB transcription factor regulates very-long-chain fatty acid biosynthesis for activation of the hypersensitive cell death response in *Arabidopsis*. *Plant Cell* 20:752–767
- Richards E, Reichardt M, Rogers S (1994) Preparation of genomic DNA from plant tissue. *Current protocols in molecular biology. Volume 1*. Edited by: Ausubel MF, Brent R, E KR, Moore DD, Seidman JG, Smith JA, Struhl K. John Wiley & Sons, Inc.;2.3.1–2.3.7.
- Romero I, Fuertes A, Benito MJ, Malpica JM, Leyva A, Paz-Ares J (1998) More than 80R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J* 14:273–284
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Schmitz G, Tillmann E, Carriero F, Fiore C, Cellinni F, Theres K (2002) The tomato *Blind* gene encodes a MYB transcription factor that controls the formation of lateral meristems. *Proc Natl Acad Sci* 99:1064–1069
- Shimizu T, Akada S, Senda M, Ishikawa R, Harada T, Niizeki M, Dube SK (1999) Enhanced expression and differential inducibility of soybean chalcone synthase genes by supplemental UV-B in dark-grown seedlings. *Plant Mol Biol* 39:785–795
- Shimizu T, Fujibe R, Senda M, Ishikawa R, Harada T, Niizeki M, Akada S (2000) Molecular cloning and characterization of a subfamily of UV-B responsive MYB genes from soybean. *Breeding Sci* 50:80–90
- Shin B, Choi G, Yi H, Yang S, Cho I, Kim J, Lee S, Paek NC, Kim JH, Song PS, Choi G (2002) *AtMYB21*, a gene encoding a flower-specific transcription factor, is regulated by COP1. *Plant J* 30:23–32
- Shin R, Burch AY, Huppert KA, Tiwari SB, Murphy AS, Guilfoyle TJ, Schachtman DP (2007) The *Arabidopsis* transcription factor MYB77 modulates auxin signal transduction. *Plant Cell* 19:2440–2453
- Stracke R, Werber M, Weisshaar B (2001) The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol* 4:447–456
- Stracke R, Ishihara H, Huep G, Barsch A, Mehrrens F, Niehaus K, Weisshaar B (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J* 50:660–677
- Takahashi M, Mukouda M, Koono K (2000) Differences in genetic structure between two Japanese beech (*Fagus crenata* Blume) stands. *Heredity* 84:103–115
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Tomaru N, Mitsutsuji T, Takahashi M, Tsumura Y, Uchida K, Ohba K (1997) Genetic diversity in *Fagus crenata* (Japanese beech); influence of the distributional shift during the late-Quaternary. *Heredity* 78:241–251
- Tominaga R, Iwata M, Okada K, Wada T (2007) Functional analysis of the epidermal-specific MYB genes CAPRICE and WERE-WOLF in *Arabidopsis*. *Plant Cell* 19:2264–2277
- Urao T, Noji M, Yamaguchi-Shinozaki K, Shinozaki K (1996) A transcriptional activation domain of ATMYB2, a drought-inducible *Arabidopsis* Myb-related protein. *Plant J* 10:1145–1148
- Vaillau F, Daniel X, Tronchet M, Montillet JL, Triantaphylides C, Roby D (2002) A R2R3-MYB gene, *AtMYB30*, acts as a positive regulator of the hypersensitive cell death program in plants in response to pathogen attack. *Proc Natl Acad Sci USA* 99:10179–10184
- Van der Ent S, Verhagen BW, Van Doorn R, Bakker D, Verlaan MG, Pel MJ, Joosten RG, Proveniers MC, Van Loon LC, Ton J, Pieterse CM (2008) *MYB72* is required in early signaling steps of rhizobacteria-induced systemic resistance in *Arabidopsis*. *Plant Physiol* 146:1293–1304
- Wang S, Wang JW, Yu N, Li CH, Luo B, Gou JY, Wang LJ, Chen XY (2004) Control of plant trichome development by a cotton fiber MYB gene. *Plant Cell* 16:2323–2334
- Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM (2009) Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. *Plant Physiol* 149:981–993
- Yoo JH, Park CY, Kim JC, Heo WD, Cheong MS, Park HC, Kim MC, Moon BC, Cho MC, Kang YH, Lee JH, Kim HS, Lee SM, Yoon HW, Lim CO, Yun DJ, Lee SY, Chung WS, Choo MJ (2005) Direct interaction of a divergent CaM isoform and the transcription factor, MYB2, enhances salt tolerance in *Arabidopsis*. *J Biol Chem* 280:3697–3706
- Zhang J, Rosenberg HF, Nei M (1998) Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc Natl Acad Sci USA* 95:3708–3713
- Zhang Y, Cao G, Qu LJ, Gu H (2009) Characterization of *Arabidopsis* MYB transcription factor gene *AtMYB17* and its possible regulation by LEAFY and AGL15. *J Genet Genomics* 36:99–107
- Zhang ZB, Zhu J, Gao JF, Wang C, Li H, Zhang HQ, Zhang S, Wang DM, Wang QX, Huang H, Xia HJ, Yang Z-N (2007) Transcription factor *AtMYB103* is required for anther development by regulating tapetum development, callose dissolution and exine formation in *Arabidopsis*. *Plant J* 52:528–538
- Zhong R, Richardson EA, Ye ZH (2007) The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*. *Plant Cell* 19:2776–2792
- Zhu J, Verslues PE, Zheng X, Lee BH, Zhan X, Manabe Y, Sokolchik I, Zhu Y, Dong CH, Zhu JK, Hasegawa PM, Bressan RA (2005) *HOS10* encodes an R2R3-type MYB transcription factor essential for cold acclimation in plants. *Proc Natl Acad Sci USA* 102:9966–9971