Theor Appl Genet (2011) 122:1561–1576
DOI 10.1007/s00122-011-1555-2

ORIGINAL PAPER

Development of PCR markers for *Tamyb10* related to *R-1*, red grain color gene in wheat

Eiko Himi · Masahiko Maekawa · Hideho Miura · Kazuhiko Noda

Received: 13 December 2010 / Accepted: 12 February 2011 / Published online: 27 February 2011

Abstract The grain color of wheat affects not only the brightness of flour, but also tolerance to preharvest sprouting. Grain color is controlled by dominant *R-1* genes located on the long arm of hexaploid wheat chromosomes 3A, 3B, and 3D (*R-A1*, *R-B1*, and *R-D1*, respectively). The red pigment of the grain coat is composed of catechin and proanthocyanidin (PA), which are synthesized via the flavonoid biosynthetic pathway. We isolated the *Tamyb10-A1*, *Tamyb10-B1*, and *Tamyb10-D1* genes, located on chromosomes 3A, 3B, and 3D, respectively. These genes encode R2R3-type MYB domain proteins, similar to TT2 of *Arabidopsis*, which controls PA synthesis in testa. In recessive *R-A1* lines, two types of *Tamyb10-A1* genes: (1) deletion of the first half of the R2-repeat of the MYB region and (2) insertion of a 2.2-kb transposon belonging to the *hAT* family. The *Tamyb10-B1* genes of recessive *R-B1* lines had 19-bp deletion, which caused a frame shift in the middle part of the open reading frame. With a transient assay using wheat coleoptiles, we revealed that the *Tamyb10* gene in the dominant *R-1* allele activated the flavonoid biosynthetic genes. We developed PCR-based markers to detect the dominant/recessive alleles of *R-A1*, *R-B1*, and *R-D1*. These markers proved to be correlated to known *R-1* genotypes of 33 varieties except for a mutant with a single nucleotide substitution. Furthermore, double-haploid (DH) lines derived from the cross between red- and white-grained lines were found to necessarily carry functional *Tamyb10* gene(s). Thus, PCR-based markers for *Tamyb10* genes are very useful to detect *R-1* alleles.

Communicated by M. Sorrells.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1555-2) contains supplementary material, which is available to authorized users.

E. Himi (✉) · M. Maekawa · K. Noda
Institute of Plant Science and Resources,
Okayama University, 2-20-1 Chuo,
Kurashiki, Okayama, Japan
e-mail: eikohimi@server.rib.okayama-u.ac.jp

H. Miura
Obihiro University of Agriculture and Veterinary Medicine,
Inada-cho, Obihiro, Hokkaido, Japan

Introduction

Preharvest sprouting (PHS) is the precocious germination of grains in the spike before harvesting. PHS in wheat induces deterioration of flour quality because starch breakdown occurs in germinated grains. Red-grained wheat varieties are usually more tolerant to PHS than white-grained wheat varieties (Flintham 2000; Warner et al. 2000; Himi et al. 2002). This association between PHS tolerance and red pigmentation of the grain is likely due to a pleiotropic effect of the genes controlling grain color (Flintham 2000; Warner et al. 2000; Himi et al. 2002). The grain color of wheat is controlled by the *R-1* (red) genes located on the distal region of the long arm of homoeologous chromosome 3 (3AL, 3BL, and 3DL). White-grained wheat varieties are homozygous for the recessive alleles *R-A1a*, *R-B1a*, and *R-D1a* (*r2*, *r3*, and *r1*, respectively, in a former notation). One or more dominant alleles, *R-A1b*, *R-B1b*, and/or *R-D1b* (*R2*, *R3*, and *R1*, respectively, in a former notation) confer red pigmentation onto grains (McIntosh et al. 1998). We previously demonstrated that the dominant *R-1* genes enhance grain dormancy using near isogenic lines carrying red grains (ANK-1A to 1D) with the genetic background of white-grained wheat (Novosibirskaya 67) and a white-grained mutant

(EMS-AUS) induced by EMS treatment of a red-grained AUS1490 line (Himi et al. 2002). In spite of the important trait of grain color in wheat breeding, genetic analysis of grain color remains inconclusive because grain color phenotypes were determined by the analysis of F_3 grains harvested from individual F_2 plants. Since grain color is determined by the maternal genotype, the phenotype of F_3 grains from F_2 plants is controlled by the F_2 genotype. Therefore, the development of genetic markers for *R-1* genotypes is required for not only genetic interests but also wheat breeding.

The pigment of red wheat grain is identified to be catechin and PAs (Miyamoto and Everson 1958; McCallum and Walker 1990). In addition, flavonols and stilbenes have been reported to be included in red grain (Matus-Cadiz et al. 2008). In barley grains, PAs consist of the oligomeric and polymeric structures of the flavan-3-ol monomers, (+)-catechin and (+)-gallocatechin, and the most abundant PAs in barley are dimeric PAs (Quinde-Axtell and Baik 2006). This suggests that PAs in wheat might also consist of oligo- and/or polymeric structures. In maize kernels, anthocyanins for purple pigment and phlobaphenes for red pigment have been characterized (Styles and Ceska 1975). *Arabidopsis* seeds accumulate flavonols and PAs (Routaboul et al. 2006). Anthocyanins, phlobaphenes, flavonols, and PAs are all synthesized through the same early flavonoid biosynthetic pathway and branched out into the individual pathway (Fig. 1). We cloned *chalcone synthase* (*CHS*), *chalcone flavanone isomerase* (*CHI*), *flavonone 3-hydroxylase* (*F3H*), and *dihydroflavonol 4-reductase* (*DFR*) genes of wheat and found that these genes were expressed predominantly in the immature grain coat of red grain but were almost completely suppressed in white-grained lines (Himi and Noda 2004; Himi et al. 2005). These results suggest that the *R-1* genes of wheat are transcription factors to regulate the flavonoid biosynthetic pathway.

Several regulatory proteins involved in flavonoid biosynthesis have been reported in various species, such as maize, petunia, snapdragon, and *Arabidopsis* (Winkel-Shirley 2001; Mol et al. 1998). Most of these transcription factors belong to two large families, MYB and bHLH, and a small family of WD40 repeats (WDR) (Stracke et al. 2001; Buck and Atchley 2003; Ramsay and Glover 2005). Anthocyanin synthesis requires the synergy of a MYB protein (such as C1/Pl in maize) and a bHLH protein (such as R/B/Lc/Sn in maize). PA synthesis in *Arabidopsis* seeds requires not only both MYB (TT2) and bHLH (TT8), but also WDR (TTG1). However, *TTG1* is expressed in both seeds and vegetative tissues, and TTG protein has other functions, such as the development of trichomes and root hairs and the accumulation of seed mucilage (Walker et al. 1999). Phlobaphene synthesis in

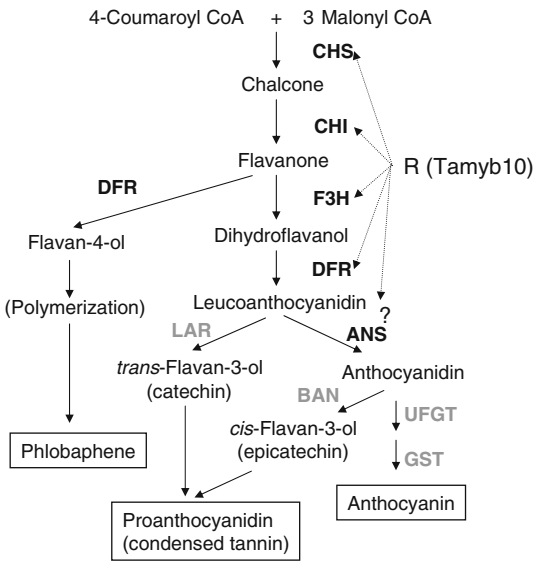

Fig. 1 The flavonoid biosynthetic pathway. Enzymes are abbreviated as follows: *CHS* chalcone synthase, *CHI* chalcone isomerase, *F3H* flavanone 3-hydroxylase, *DFR* dihydroflavonol 4-reductase, *ANS* anthocyanidin synthase, *UFGT* UDPG-flavonoid glucosyl transferase, *BAN* banyuls, *LAR* leucoanthocyanidin reductase, *GST* glutathione *S*-transferase. The pathway is demonstrated according to Winkel-Shirley (2001) and Xie et al. (2004). Wheat genes for the enzymes shown in *gray* have not yet been cloned

maize kernel is regulated by P, a MYB protein, alone (Grotewold et al. 1994). Since each wheat *R-1* gene appears to regulate flavonoid biosynthesis and the accumulation of flavonoid pigments in grain coat, we presumed that *R-1* genes may encode MYB-type proteins, such as P of maize, and cloned 14 *MYB* genes from wheat. One of them, *Tamyb10*s, is located on the distal region of the long arm of chromosomes 3A, 3B, and 3D (Himi and Noda 2005).

In this paper, we demonstrate the variation of sequences of *Tamyb10* genes from red- and white-grained wheat varieties, and *Tamyb10* genotypes are consistent with the *R-1* gene genotypes of the varieties, which were previously clarified from genetic analysis by crossing. The deduced Tamyb10 amino acid sequences have the same motif with a TT2 of *Arabidopsis* and OsMYB3 of rice. When the *Tamyb10-A1* gene of red-grained AUS1490 and the *Tamyb10-D1* genes of red-grained Chinese Spring (CS) with the CaMV35S promoter were delivered to a colorless coleoptile of CS, anthocyanin accumulation was observed. Thus, *Tamyb10* is likely to be a strong candidate for the *R-1* gene of wheat, which regulates wheat grain color. Furthermore, a novel transposon belonging to the *hAT* family, *Genome Surfing Trader* (*GeST*), was found in the *Tamyb10-A1* gene of some *R-A1a* lines. From all of this information, we developed PCR-based markers to detect an *R-1* genotype easily at an early stage.

Materials and methods

Plant materials

Wheat (*Triticum aestivum*) varieties or lines used in this study are listed in Table 1. DH populations used in this study were derived from F_1 plants of the cross between red-grained Zenkoji Komugi (*R-A1a*/*R-B1b*/*R-D1b*) and white-grained Tamaizumi (*R-A1a*/*R-B1a*/*R-D1a*) consisting of 142 individuals. They were grown under a semi-transparent plastic roof house at Institute of Plant Science and Resources, Kurashiki, Japan. The spikes were tagged at anthesis and harvested 5 days post-anthesis. Grains were collected from the primary and secondary florets of the central spikelets of the spikes. Leaves and roots were collected from 3-day-old seedlings germinated on plastic plates with 2 pieces of filter paper. For DNA and RNA extraction, developing grains and seedlings grown at 20°C under 12 h of UV light (about 100 mol m^{-2} s^{-1}, UV lamp) or dark conditions were used.

Genomic DNA, RNA, and cDNA preparation

DNA was isolated from 1 g of 10-day-old seedlings according to Murray and Thompson (1980). Total RNA was extracted from about 0.5 g of grains and 1 g of leaves and roots of 3-day-old seedlings by the SDS-phenol method (Himi and Noda 2004). Poly (A)+ RNA was isolated from 10 mg of total RNA with an mRNA isolation kit (Roche Diagnostics) according to the instructions from the supplier. cDNA was obtained by the reverse-transcription reaction using SuperScript II (Invitrogen).

Isolation of *Tamyb10-A1*, *Tamyb10-B1*, and *Tamyb10-D1*

All primers used in this study are listed in Table 2. The positions of primers are shown in Suppl Fig. S1. Rapid amplification of cDNA 3′ ends (3′ RACE) was performed to isolate a 3′ region of *Tamyb10* using mRNA of CS grains of 5 DPA. The primers of P1LP and P2LP were designed on the basis of the MYB consensus region of the *ZmP* gene (accession number Z11879). The 3′ region of Tamyb10 cDNA was amplified with P1LP and 3′ adapter primers, and nested PCR was carried out with P2LP and 3′ adapter primers. Cloning of PCR fragments was done with pGEM-T vector systems (Promega). DNA sequences were determined with the PRISM 310 ABI DNA sequencing system (Applied Biosystems).

Three primers, Tamyb10-LP1, Tamyb10-RP1, and Tamyb10-RP2, were designed for the 3′ RACE product, except for the MYB consensus region. Tamyb10-LP1 and Tamyb10-RP1 primers could recognize and amplify a partial sequence of *Tamyb10-A1* from the PCR with genomic DNA of N3A3TD, N3BT3D, and N3DT3A (Suppl Fig. S2a). To obtain *Tamyb10-B1* and *Tamyb10-D1*, PCR was performed using an N3AT3D line with the sets of P1LP and Tamyb10-RP1 and P1LP and Tamyb10-RP2 in low stringency. Two different sequences were obtained, and primers that recognize each sequence were designed (Tamyb10-LP2 and Tamyb10-LP3). These primers were used for 3′ RACE, and each product was sequenced. Tamyb10-RP3 was designed in the latter 3′ RACE products. The primer set of Tamyb10-LP2 and Tamyb10-RP1 could recognize *Tamyb10-B1*, and the primer set of Tamyb10-LP3 and Tamyb10-RP3 could recognize *Tamyb10-D1* using N3AT3D, N3BT3D, and N3DT3A lines (Suppl Fig. S2b, c). The upstream regions of each *Tamyb10* gene were isolated by the inverse PCR method. Genomic DNA of CS was treated with *Pvu*II and self-ligated. Then, PCR using Tamyb10-LP4 and Tamyb10-RP4 was performed, and the products were used for nested PCR with 3 sets of primers (Tamyb10-LP1 and Tamyb10-RP4, Tamyb10-LP2 and Tamyb10-RP4, and Tamyb10-LP3 and Tamyb10-RP4). The Tamyb10-LP5 and Tamyb10-RP5 primers, which were designed in the region including the start codon and the stop codon, respectively, were used to obtain full sequences of the *Tamyb10* gene.

Expression analysis of the *Tamyb10* gene

RT-PCR for examining the level of *Tamyb10* gene expression was conducted using Tamyb10-LP4 and Tamyb10-RP1 primers. The primers for *CHS*, *CHI*, *F3H*, *DFR*, and *ANS* genes are listed in Table 2. Primers ubi-LP and ubi-RP for the *ubiquitin* gene and TaActin LP and TaActin RP for the *actin* gene were used as internal controls.

PCR genotyping for *Tamyb10-A1*, *B1*, and *D1*

The primers and PCR conditions are described in Table 2 and Suppl Fig. S3a–e. *Tamyb10-A1* genes are classified into three groups (functional *Tamyb10-A1* of Norin 61 type (*R-A1b*), non-functional *Tamyb10-A1* of CS type (*R-A1a*), and non-functional *Tamyb10-A1* of Norin 17 type (*R-A1a*)). Therefore, three PCR should be required to distinguish each *Tamyb10-A* gene.

Transient assay

A transient vector, pBI221 (AF502128), which has the beta-glucuronidase gene driven by the CaMV 35S promoter, was used for the transient assay. We inserted *Tamyb10-D1* of CS, *Tamyb10-A1* of AUS1490, or *Tamyb10-A1* of EMS-AUS genes instead of the beta-glucuronidase gene of pBI221 and the designated pBI-myb10D (CS), pBI-myb10A (AUS), or pBI-myb10A (EMS),

Table 1 Genotypes of *R-1* gene in red-or white-grained lines or varieties

	R-A1	*R-B1*	*R-D1*	Former notation	Grain color
CS and its aneuploid lines[a]					
CS	*a*	*a*	*b*	*R1r2r3*	*Red*
N3AT3D	None	*a*	*b*	*R1r3*	*Red*
N3BT3D	*a*	None	*b*	*R1r2*	*Red*
N3DT3A	*a*	*a*	None	*r2r3*	White
Novosibirskaya 67 and its near isogenic lines[b]					
Novosibirskaya 67 (NS67)	*a*	*a*	*a*	*r1r2r3*	White
ANK-1A	*a*	*a*	*b*	*R1r2r3*	*Red*
ANK-1B	*a*	*b*	*a*	*r1r2R3*	*Red*
ANK-1C	*a*	*a*	*b*	*R1r2r3*	*Red*
ANK-1D	*b*	*a*	*a*	*r1R2r3*	*Red*
ANK-1E	*a*	*b*	*a*	*r1r2R3*	*Red*
Varieties/lines					
Norin61	*b*	*b*	*b*	*R1R2R3*	*Red*[c]
RL4137	*b*	*b*	*b*	*R1R2R3*	*Red*[c]
Fukuho Komugi	*b*	*b*	*a*	*r1R2R3*	*Red*[c]
NorinIO	*b*	*a*	*b*	*R1R2r3*	*Red*[c]
Asakaze Komugi	*b*	*a*	*b*	*R1R2r3*	*Red*[c]
Norin66	*b*	*a*	*a*	*r1R2r3*	*Red*[c]
Fukuwase Komugi	*b*	*a*	*a*	*r1R2r3*	*Red*[c]
AUS1490	*b*	*a*	*a*	*r1R2r3*	*Red*[d]
EMS-AUS	*a*	*a*	*a*	*r1r2r3*	White[d]
Kitakami Komugi	*a*	*b*	*b*	*R1r2R3*	*Red*[c]
Zenkoji Komugi	*a*	*b*	*b*	*R1r2R3*	*Red*[e]
Norin17	*a*	*a*	*a*	*r1r2r3*	White
Tamaizumi	*a*	*a*	*a*	*r1r2r3*	White
Ackarma	*a*	*a*	*a*	*r1r2r3*	White
SZF	*a*	*a*	*a*	*r1r2r3*	White
Cornell 595	*a*	*a*	*a*	*r1r2r3*	White
8019R1	*a*	*a*	*a*	*r1r2r3*	White
8021V2	*a*	*a*	*a*	*r1r2r3*	White
AUS1408	*a*	*a*	*a*	*r1r2r3*	White
BL1496	*a*	*a*	*a*	*r1r2r3*	White
Cadoux	*a*	*a*	*a*	*r1r2r3*	White
Clark's cream	*a*	*a*	*a*	*r1r2r3*	White
Gaines	*a*	*a*	*a*	*r1r2r3*	White
Hakei91-64	*a*	*a*	*a*	*r1r2r3*	White
Prina	*a*	*a*	*a*	*r1r2r3*	White
Ryubaku7	*a*	*a*	*a*	*r1r2r3*	White

White-grained wheats are homozygous for the recessive alleles *R-A1a*, *R-B1a*, and *R-D1a*; red-grained wheats carry one or more of the dominant alleles *R-A1b*, *R-B1b*, and *R-DIb*, according to Flintham (2000)

[a] Sears (1944), Allan and Vogel (1965), Metzger and Silbaugh (1970)

[b] Koval (1997)

[c] Kaneko et al. (1994)

[d] Mares et al. (2005)

[e] Miura et al. (2002)

Table 2 List of primers used in this study

Primer name	Primer sequence (5′–3′)
Primer sequences for *Tamyb10* genes	
P1LP	ccggtggatcaactaccttcg
P2LP	cggccgaacagacaacgacat
3′Adaptor	ggccacgcgtcgactagtac
Tamyb10-LP1	ctgagcaagaggatgctgc
Tamyb10-RP1	gatgccctccagatcaaggt
Tamyb10-LP2	agcaagaggaacctgcagtc
Tamyb10-RP2	cgtgaaaaccctctgctacc
Tamyb10-LP3	cgctgagcaagaggaacca
Tamyb10-RP3	aggcacaccagcttatttgg
Tamyb10-LP4	cgaacagacaatgagatcaagaa
Tamyb10-RP4	cctgacgatgagctcctctt
Tamyb10-LP5	cgggatccatggggaggaaaccatgctg
Tamyb10-RP5	cggatatctcaaagccacgccaactcca
Tamyb10-LP6	ctatgtggatggccttggat
Tamyb10-RP6	ctaccagctcgtttgggaag
Tamyb10-LP7	tttcaatcgagtgggcataa
Tamyb10-RP7	tgttatcacatgctgatcctga
Tamyb10-LP8	tccctacatgggagacagaga
Tamyb10-LP9	taggccaacaccttctaaacg
Primer sequences for expression analysis	
ubi-LP	atttgtg aag accctcaccg
ubi-RP	caccaagtgaagggtggact
TaActin LP	gagggatacacgcttcctca
TaActin RP	gaaagtgctaagagaggccaaa
CHS-LP	atcacccacctcgtcttctg
CHS-RP	aggaggtggaaggtgagtcc
TaCHI1-LP	gcagtactcggacaaggtga
TaCHI2-RP	gttcgttcacaccgaaacc
F3H1-LP	cctacttctctcgtacccggtg
F3H2-RP	gaacgtcgcgatcgacag
A1-LP	ctcatggctcgtcaggaag
DFR1-RP	tcttgggagtcgaagtccat
TaANSsense	gtctccgcgctctccttc
TaANSantisense	gcccgttgctgaggatct

respectively. These genes were amplified with the Tamyb10-LP5 and Tamyb10-RP5 primers using cDNA of 5 DPA of CS, AUS1490, and EMS-AUS grains, respectively. Grains of CS were imbibed at 25°C for 48 h under dark conditions. pBI-myb10D (CS), pBI-myb10A (AUS), pBI-myb10A (EMS), and pBI221 (control vector) were delivered into CS coleoptiles by particle bombardment according to our previous report (Ahmed et al. 2003). Pigmentation of coleoptiles was observed after 24 h at 25°C under dark conditions.

A transient assay using pBI-myb10D (CS) was also performed with coleoptiles of white-grained varieties (listed in Table 4).

Results

Isolation of *Tamyb10* genes

To complete the full sequence of *Tamyb10* genes expressed predominantly in immature red grains (Himi and Noda 2005), we then cloned and sequenced the 3′ region of *Tamyb10* in cDNA derived from the immature grains of wheat cv. Chinese Spring (CS; *R-A1a/R-B1a/R-D1b*) by the 3′RACE method. We also amplified the partial sequences of *Tamyb10* from the genomic DNA of normal CS and three chromosome addition-deletion lines of CS, Nulli3ATetra3D (N3AT3D), Nulli3BTetra3D (N3BT3D), and Nulli3DTetra3A (N3DT3A), which lacked chromosomes 3A, 3B, and 3D, respectively. Thus, we identified three *Tamyb10* genes located on chromosomes 3A, 3B, and 3D (Fig. 2; Suppl Fig. S2a–c). Each fragment included a MYB region and characteristic intron sequences. The 5′ regions of the *Tamyb10* genes on chromosomes 3A, 3B, and 3D of CS were obtained by the inverse PCR method. The full genomic sequences of the *Tamyb10* genes on chromosomes 3A, 3B, and 3D of CS were then cloned using primers specific to each *Tamyb10* gene. We designated the *Tamyb10* genes on chromosomes 3A, 3B, and 3D as *Tamyb10-A1* (Genbank no. AB191458), *Tamyb10-B1* (AB191459), and *Tamyb10-D1* (AB191460), respectively.

Furthermore, to elucidate the relationship between *Tamyb10* and dominant *R-1* genotype, *Tamyb10-A1* of AUS1490 carrying *R-A1b* and *Tamyb10-B1* of Norin 61 carrying *R-B1b* were isolated (Table 1). These sequences were isolated from both genomic DNA and cDNA (derived from the immature grains).

Thus, *Tamyb10-A1*, *B1*, and *D1* derived from each dominant allele of *R-1* gene have 3 exons and 2 introns (Fig. 2a) and encode 259, 268, and 265 amino acid residues, respectively (Suppl Fig. S4). The deduced amino acid sequences of *Tamyb10-A1* showed high identity to those of *Tamyb10-B1* and *D1* (90 and 91%, respectively), and the deduced amino acid sequence of *Tamyb10-B1* also showed high identity to that of *Tamyb10-D1* (91%) (Suppl Fig. S4).

A phylogenetic tree was constructed using the predicted amino acid sequences of the R2R3 MYB domains of *Tamyb10* genes and flavonoid regulatory R2R3 MYB proteins from other species (Fig. 3). Tamyb10s, OsMYB3, and SbMYB1 are most closely related to the PA-clade 2 rather than to other flavonoid regulatory MYB subgroups, such as PA-clade 1 and other anthocyanin- and phlobaphene regulators. While the functions of OsMYB3 and SbMYB1 have not been identified, these MYB proteins and Tamyb10s seem likely to act as PA regulators.

Furthermore, the amino acid sequences encoded by *Tamyb10*s share a conserved sequence (IRTKAL/IRC) between *Arabidopsis* TT2 (Nesi et al. 2001) and rice

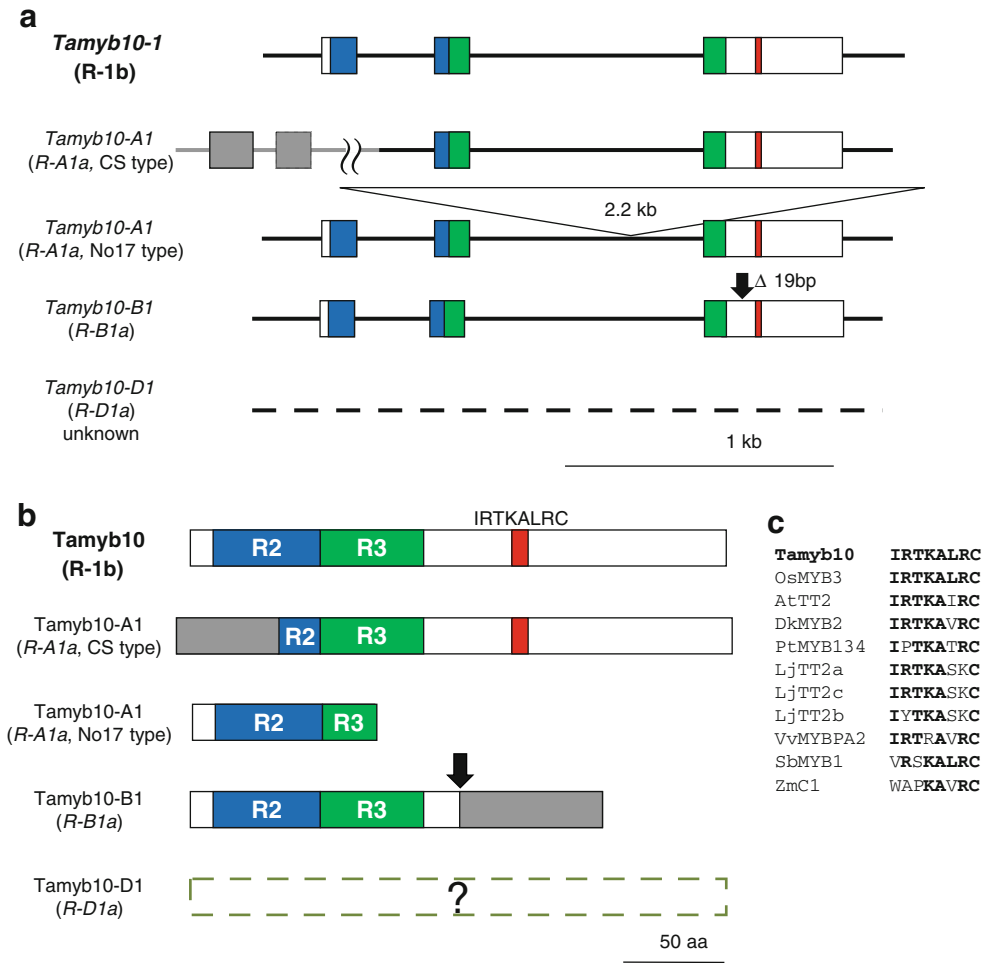

Fig. 2 Structure of *Tamyb10* genes and putative Tamyb10 proteins. **a** Genomic organization of the *Tamyb10-1* genes. *Boxes* indicate exons, and *lines* indicate introns and untranslated regions. Functional *Tamyb10-A1*, *Tamyb10-B1*, and *Tamyb10-D1* from the lines of *R-A1b*, *R-B1b*, and *R-D1b* alleles, respectively, have R2 (*blue box*) and R3 (*green box*) repeats of the MYB consensus region and conserved sequences (IRTKAL/IRC, *red box*) among Tamyb10, *Arabidopsis* TT2, and rice OsMYB3. *Tamyb10-A1* gene (CS type) of recessive *R-A1a* alleles (such as CS) was rearranged in the upper region to an unknown sequence (*grey lines and boxes*). The *hAT* family-like 2.2-kb sequence (*GeST*) was inserted into a second intron of the *Tamyb10-A1* gene (Norin 17 type) of the recessive *R-A1a* alleles (such as Norin 17). *Tamyb10-B1* gene of the recessive *R-B1a* alleles had a 19-bp deletion in the third exon (*black arrow*). *Tamyb10-D1* gene is unidentified. *Bar* 1 kb. **b** Functional Tamyb10 encoded by *Tamyb10-A1*, *Tamyb10-B1*, and *Tamyb10-D1* from the lines of *R-A1b*, *R-B1b*, and *R-D1b* alleles, respectively, has R2 (*blue box*) and R3 (*green box*) repeats of the MYB consensus region. The IRTKAL/IRC motif is represented by a *red box*. Tamyb10-A1 of CS (*R-A1a*) altered the first half of the R2 repeat (*grey box*). Tamyb10-A1 of Norin 17 (*R-A1a*) was truncated to lack the last half of the R3 repeat to C-terminus region. Tamyb10-B1 of recessive *R-B1a* alleles had a 19-bp deletion in the middle part (*black arrow*) and was frame-shifted to different amino acid sequences after the deletion (*grey box*). The Tamyb10-D1 sequence of the *R-D1a* allele is still unknown. *Bar* 50 amino acids. **c** Region of conserved sequences (IRTKAL/IRC) of Tamyb10 (wheat), OsMYB3 (rice), AtTT2 (*Arabidopsis*), DkMYB2 (persimmon), PtMYB134 (quaking aspen), LjTT2a-c (*Lotus japonicas*), VvMYBPA2 (grape), SbMYB1 (sorghum), and ZmC1 (maize)

OsMYB3 (Suzuki et al. 1997) in the middle of the sequences (Fig. 2a–c; Suppl Fig. S4). This motif is also found in DkMYB2 (Akagi et al. 2010), PtMYB134 (Mellway et al. 2009), LjTT2s (Yoshida et al. 2008), VvMYBPA2 (Terrier et al. 2009), and SbMYB1 (accession number: GU479928) (Fig. 2c). TT2 is necessary for the expression of late flavonoid biosynthetic genes (PA pathway genes), such as *DFR*, *anthocyanidin synthase* (*ANS*), and *banylus* (*BAN*), in immature seeds and for PA accumulation in seed testa (Fig. 1). DkMYB2, PtMYB134, LjTT2s, and VvMYBPA2 were also reported as regulators to accumulate PA. While more than 100 R2R3 MYB genes have been identified in *Arabidopsis*, no significant similarity was found within the C-terminal regions between TT2 and any other MYB protein (Nesi et al. 2001). On the other hand, DkMYB4 (Akagi et al. 2009) and VvMYBPA1 (Bogs et al. 2007) were also reported as PA regulators, but these proteins do not have this motif and can be classified into another clade (PA-clade 1) rather than the *Arabidopsis* TT2 group (PA-clade 2) (Akagi et al. 2010).

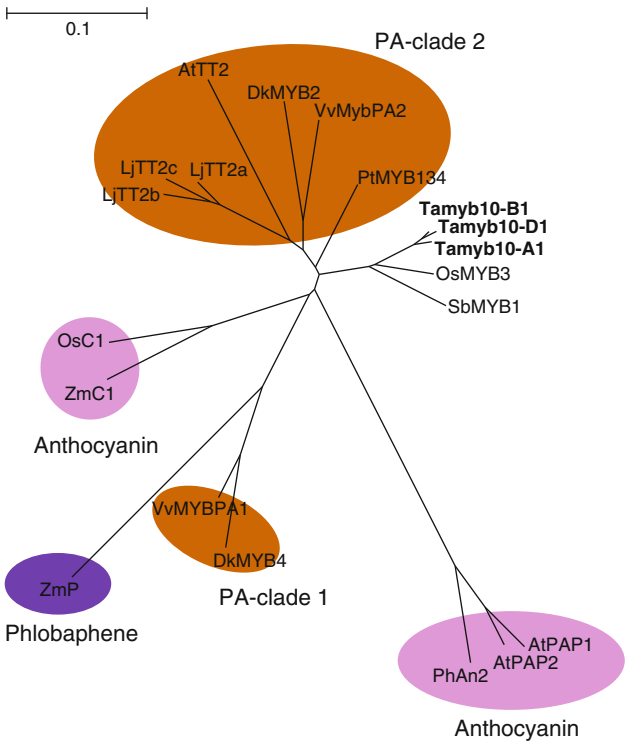

Fig. 3 Phylogenetic analysis showing plant MYB transcription factors. The tree was constructed from the ClustalW alignment using the neighbor-joining method. The *scale bar* represents 0.1 substitutions per site. The GenBank accession numbers of the MYB proteins are as follows: Tamyb10-A1, B1, and D1 (wheat, AB599721, AB599722, and AB191460), OsMYB3 (rice, BAA23339), SbMYB1 (sorghum, ADD18214), AtPAP1 (*Arabidopsis*, AAG42001), AtPAP2 (*Arabidopsis*, AAG42002), DkMYB4 (persimmon, BAI49721), VvMYBPA1 (grape, CAJ90831), ZmP (maize, AAC49394), ZmC1 (maize, P10290), OsC1 (rice, BAD04030), LjTT2a-c (*Lotus japonicas*, BAG12893-BAG12895), AtTT2 (*Arabidopsis*, NP_198405), DkMYB2 (persimmon, BAI49719), VvMYBPA2 (grape, ACK56131), and PtMYB134 (quaking aspen, ACR83705)

The N-terminal amino acid sequence encoded by *Tamyb10-A1* of CS (*R-A1a*) was quite different from that of the *Tamyb10-A1* genes of AUS1490 (Fig. 2a, b; Suppl Fig. S4). The Tamyb10-A1 protein of CS appears to lack the function to bind DNA because of the loss of the first half of R2 repeat of MYB domain. On the other hand, another *R-A1* recessive line, Norin 17 (*R-A1a*), encodes a similar N-terminal amino acid sequence to Tamyb10-A1 of AUS1490 (*R-A1b*) but has a 2.2-kb insertion into the second intron (Fig. 2a). This insertion may cause incomplete transcription, since the expression of *Tamyb10-A1* in Norin 17 was not detected by RT-PCR (Himi and Noda 2005). While the Tamyb10-B1 of CS has conserved R2R3 repeats, a 19-bp deletion of the CCG repeat region on the downstream of R2R3 repeats caused a frame shift, resulting in different amino acid sequences from that of Tamyb10-B1 of Norin 61 (Fig. 2a, b; Suppl Fig. S4). On the other hand, no *Tamyb10-D1* of other *R-D1a* varieties was isolated using specific primers for *Tamyb10-D1* or common primers for *Tamyb10*s, suggesting that *Tamyb10-D1* genes of *R-D1a* varieties might be deleted.

Linkage between *R-1* genotype and *Tamyb10* genotype

We investigated *Tamyb10* genotypes with 33 wheat varieties with known *R-1* genotypes using *Tamyb10-A1*-, *B1*-, and *D1*-specific primers. To detect 3 types (the functional type, non-functional CS type, and non-functional Norin 17 type) of *Tamyb10-A1* (Fig. 2a, b), we designed 3 sets of primers for: (1) an upstream region of the functional type, (2) an upstream region of the non-functional-CS type, and (3) a second intron with/without a 2.2-kb insertion (Suppl Fig. S3). All *R-A1b* varieties (Norin 61, RL4137, Fukuho Komugi, Norin 10, Asakaze Komugi, Norin 66, Fukuwase Komugi, and AUS1490) showed a 665-bp fragment with primer set (1) and a 565-bp fragment with primer set (3) but not with primer set (2) (Fig. 4a, b). These results showed that these lines have the same type of sequences that encode the R2R3 repeat and the IRTKAL/IRC motif. Although three *R-A1a* varieties (Norin 17, Kitakami Komugi, and AUS1408) also showed the same amplified pattern as *R-A1b* varieties with primer sets (1) and (2), 2,750-bp fragments were amplified with the primer set (3) (Fig. 4a, b). This result suggested that these three lines lack the lower region of Tamyb10-A1, where the last half of the R3 repeat and the IRTKAL/IRC motif exist because of the 2.2-kb insertion in the second intron (Fig. 2b). Other *R-A1a* varieties (CS, Zenkoji Komugi, Tamaizumi, Ackarma, SZF, Cornell 595, 8019R1, 8021V2, BL1496, Cadoux, Clark's cream, Gaines, Hakei 91-64, Prina, and Ryubaku 7) showed amplified fragments with primer sets (2) and (3) but not with primer set (1) (Fig. 4a, b). Therefore, Tamyb10-A1 of these lines lost the function, since the first half of R2 repeat is deleted though IRTKAL/IRC motif exists. Furthermore, *Tamyb10-A1* was examined in white-grained EMS-AUS with *R-A1a* induced by EMS treatment of red-grained AUS1490 with *R-A1b* (Mares et al. 2005). The PCR amplified patterns were exactly the same as those AUS1490 using primer sets (1), (2), and (3) for *Tamyb10-A1*. However, we found a single nucleotide substitution of guanine to adenine at position 52, resulting in a single amino acid substitution of glycine (G) to glutamic acid (E) at position 15 (Suppl Fig. S4). These results indicated that the PCR-fragment patterns of *Tamyb10-A1* using 3 specific sets of primers and a single nucleotide change of EMS-AUS are completely correlated to the *R-A1* genotypes in 33 varieties/lines. Therefore, these sets of primers for *Tamyb10-A1* are useful to distinguishing the *R-A1* genotype.

The linkage between the *Tamyb10-B1* and *R-B1* genotypes was also examined using presence/absence detection

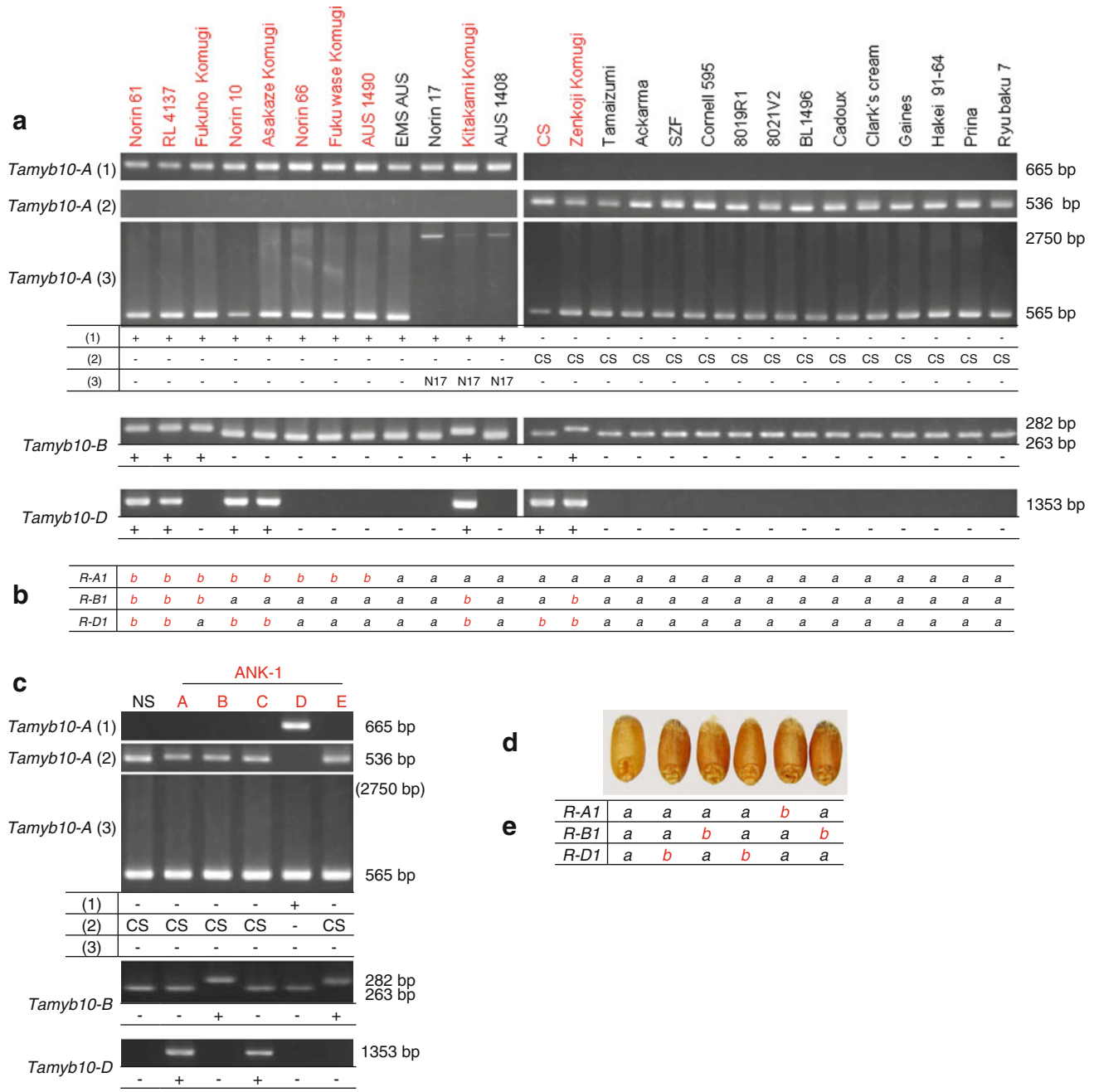

Fig. 4 Amplification profile of *Tamyb10-A1*, *B1*, and *D1* genes with sets of Tamyb10-A (1), (2), and (3), respectively, and Tamyb10-B and Tamyb10-D-specific primers with wheat varieties/lines of the known *R-1* genotype. **a**, **c** PCR patterns of each *Tamyb* gene. In tables under gel image, the + indicates PCR amplification with primer sets for *Tamyb10-A1* (1) and *Tamyb10-D1* and PCR amplification of a larger fragment with primer sets for *Tamyb10-B1*. *CS* indicates PCR amplification with primer sets for *Tamyb10-A1* (2) and *N17* indicates PCR amplification of a larger fragment with primer sets for *Tamyb10-A1* (3). The wheat varieties/lines used in this study are listed in Table 1, and the varieties/lines written in *red letters* indicate red-grained varieties/lines. The sequences and positions of the primers and PCR conditions are shown in Table 2 and Suppl Fig. S3, respectively. The sizes (bp) of amplified fragments are shown on the right-hand side. **b**, **e** *R-1* genotype of each wheat line. Dominant alleles (*R-A1b*, *R-B1b*, and *R-D1b*) are listed as *b* in *red*, and recessive alleles (*R-A1a*, *R-B1a*, and *R-D1a*) are listed as *a*. **d** grain color of Novosibirskaya 67 (white-grained variety) and its near isogenic lines, ANK-1A to 1E (red-grained lines)

primers for 19-bp deletion in exon 2 of *Tamyb10-B1* (Suppl Fig. S3). 282-bp fragments were amplified in all *R-B1b* varieties (Norin 61, RL4137, Fukuho Komugi, Kitakami Komugi, and Zenkoji Komugi). On the other hand, a 263-bp fragment (19-bp deleted fragments) was found in all *R-B1a* varieties (Norin 10, Asakaze Komugi, Norin 66,

Fukuwase Komugi, AUS1490, EMS-AUS, Norin 17, AUS1408, CS, Tamaizumi, Ackarma SZF, Cornell 595, 8019R1, 8021V2, BL1496, Cadoux, Clark's cream, Gaines, Hakei 91-64, Prina, and Ryubaku 7) (Fig. 4a, b).

Furthermore, primers for *Tamyb10-D1* derived from CS amplified 1,353-bp fragment in all *R-D1b* varieties (Norin 61, RL4137, Norin 10, Asakaze Komigi, Kitakami Komugi, CS, and Zenkoji Komugi) but no amplification was found in any of the *R-D1a* varieties (Fukuho Komugi, Norin 66, Fukuwase Komugi, AUS1490, EMS-AUS, Norin 17, AUS1408, Tamaizumi, Ackarma SZF, Cornell 595, 8019R1, 8021V2, BL1496, Cadoux, Clark's cream, Gaines, Hakei 91-64, Prina, and Ryubaku 7) (Fig. 4a, b).

These correlations between the *R-1* genotypes and PCR-fragment patterns of *Tamyb10* were also confirmed in white-grained Novosibirskaya 67 and its near isogenic lines, ANK-1A to 1E, carrying only one allele of *R-1b* each. Whereas Novosibirskaya 67 has the non-functional CS type of *Tamyb10-A1*, 19-bp deleted *Tamyb10-B1*, and non-amplified *Tamyb10-D1*, introduction of one of each dominant allele of *R-1b* into Novosibirskaya 67 was coincident with the introduction of each functional type of *Tamyb10*, as shown in Fig. 4c–e.

PCR genotyping of 142 DH lines produced from the cross between Zenkoji Komugi (*R-A1a/R-B1b/R-D1b*) and Tamaizumi (*R-A1a/R-B1a/R-D1a*) was conducted. All 36 white-grained lines showed 19-bp deleted amplification (263 bp) with a set of primers for *Tamyb10-B1* and no amplification with a set of primers for *Tamyb10-D1* (Table 3). On the other hand, red-grained 106 lines were grouped into 3 types: (1) 282-bp amplification with *Tamyb10-B1* primers and no amplification with *Tamyb10-D1* primers (31 lines), (2) 263-bp amplification with *Tamyb10-B1* primers and 1,353-bp amplification with *Tamyb10-D1* primers (36 lines), and (3) 282-bp amplification with *Tamyb10-B1* primers and 1,353-bp amplification with *Tamyb10-D1* primers (39 lines). Although none of the *R-1* genotypes of any of the red-grained lines could be determined, it was clear that at least one functional type of *Tamyb10* produced red color in the grain. These results showed that the genome-specific primer sets could be useful to detect each allele of the *R-1* gene in wheat breeding programs.

Table 3 Cosegregation of genotype of *Tamyb10-B1* and *D1*, and grain color in doubled haploid (DH) lines derived from the cross between Zenkoji Komugi and Tamaizumi

Amplification pattern of *Tamyb10-B1*[a]	+	+	−	−
Amplification pattern of *Tamyb10-D1*[b]	+	−	+	−
Estimated genotype of *R-B1*	*b*	*b*	*a*	*a*
Estimated genotype of *R-D1*	*b*	*a*	*b*	*a*
Grain color	Red	Red	Red	White
No. of lines	39	31	36	36

χ^2 for *R-B1b/R-D1b* (1):*R-B1b/R-D1a* (1):*R-B1a/R-D1b* (1):*R-B1a/R-D1a* (1) = 0.81; χ^2 for white (1):red (3) = 0.92

[a] + indicates 282-bp amplification and − indicates 263-bp (19-bp deleted) amplification with a set of primers for *Tamyb10-B1*

[b] + indicates 1,353-bp amplification and − indicates no amplification with a set of primers for *Tamyb10-D1*

An inserted 2.2-kb sequence, *Genome Surfing Trader* (*GeST*), into the second intron in *Tamyb10-A1* is a transposon-like sequence belonging to the *hAT* family

The inserted 2.2-kb sequences and the insertion site into the second intron of *Tamyb10-A1* in Norin 17, AUS1408, and Kitakami Komugi are identical. The inserted sequence was revealed to have 17-bp terminal inverted repeats (TIRs) on both ends and to be flanked by 8-bp target site duplications (TSDs) (Fig. 5a, b). Among the transposases of *hAT*, BED zinc finger motif near N-terminus and *hAT* dimerization motif at C-terminus are highly conserved (Huang et al. 2009) (Fig. 5c). Furthermore, a BLASTX search revealed that the 2.2-kb sequence possesses a putative BED zinc finger motif at N-terminus that the *hAT* family transposons commonly carry (Fig. 5b). These characteristics suggested that the inserted sequence likely belongs to the *hAT*-family transposon (Kempken and Windhofer 2001). However, the putative amino acid sequence of the 2.2-kb sequence is short, and *hAT* dimerization motif, which is highly conserved in the transposase of the *hAT* family transposons, was not detected in the sequence (Fig. 5b). From these results, the 2.2-kb sequence, named *Genome Surfing Trader* (*GeST*), may be changed from the original active autonomous form to the 2.2-kb sequence as a non-autonomous form.

Since *GeST*s of Norin 17, AUS1408, and Kitakami Komugi are the same and are inserted into the same position (Fig. 5a), a progenitor of these varieties is expected to carry the same *GeST* in the same insertion site. AUS1408 was originated from the Transvaal region of South Africa (Mares et al. 2009). Norin 17 and Kitakami Komugi were bred in Japan and registered to the Ministry of Agriculture, Forestry, and Fisheries (MAFF) in 1936 and 1959, respectively (The NIAS Genebank, http://www.gene.affrc.go.jp/index_en.php).

Since the highly strong linkage between the *R-1* genotype and *Tamyb10* genotype is revealed, the *R-1* genotypes of progenitors were estimated through PCR amplification patterns using *Tamyb10*-detection-specific primer sets (Fig. 6a, b). Although some lines were unavailable from the NIAS Genebank, we were able to estimate the *R-1* genotype(s) from the parents and descendants. For example, the *R-1* genotype of F5-31 line was estimated to be *R-A1b/R-B1b/R-D1b* (*R1R2R3* in a former style) because both Shiro Daruma and Velvet as its parents carry all

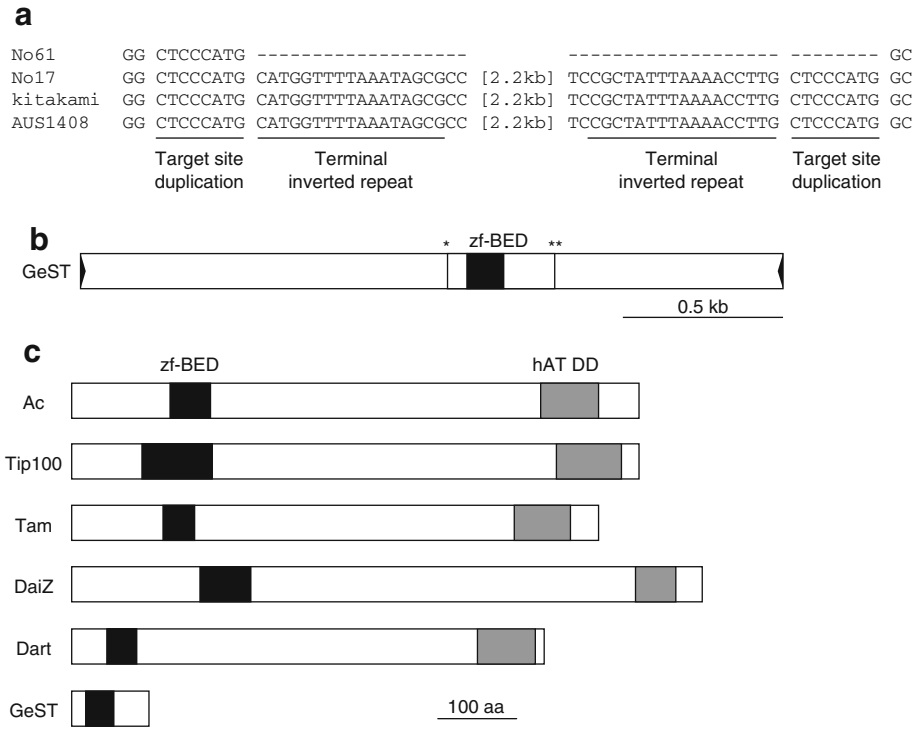

Fig. 5 Structures of *GeST*. **a** Sequence of the *GeST* insertion site. The sequences of target site duplication (TSD) and terminal inverted repeat (TIR) are *underlined*. **b** A structure of *GeST*. A region with similarity to the zinc finger-BED motif (zf-BED) is marked with a *black box*. Putative initiation and termination sites of coding region are shown with *single and double asterisks*, respectively. The *black triangles* at both ends represent TIRs. *Bar* 0.5 kb. **c** Structures of putative transposase proteins of *hAT* family in plant. The *black boxes* and *grey boxes* represent zf-BED and *hAT* dimerization motifs, respectively. The GenBank accession numbers of the proteins are as follows: Ac (maize, P08770), Tip100 (common morning glory, Q9ZWT4), Tam (snapdragon, CAA38906), DaiZ (rice, ACN38703), and Dart (rice, BAI39457)

dominant alleles of *R-1*. The *R-1* genotype of Kitakanto 5 was also estimated to be *R-A1a* (Norin 17 type) /*R-B1b* (*r2R3*) since both parents have *R-B1b* (*R3*), and its progeny, Tohoku 71, has *R-A1a* (Norin 17 type) (*r2*) and Norin 10 as the parent of Tohoku 71 has *R-A1b/R-B1a/R-D1b* (*R1R2r3*). From these results, Goshu 13, as the common progenitor of Kitakami Komugi and Norin 17, was presumably the origin of *GeST*. The *R-1* genotype of Goshu 13 was not surveyed because the seeds are not preserved in the NIAS Genebank. Interestingly, the *R-1* genotype of Goshu Komugi, which is available from NIAS, was *R-A1a* (Norin 17 type)/*R-B1a*/*R-D1a*, and "Goshu" is synonymous with "Australia". Although there is no evidence to demonstrate the relation between the progenitor of Goshu 13/Goshu Komugi and AUS1408, these varieties possibly have the same progenitor, which has the *GeST* insertion.

Activation of flavonoid biosynthetic genes by Tamyb10

Since the red pigment of wheat grain was identified as catechin and PAs (Miyamoto and Everson 1958; McCallum and Walker 1990), biosynthesis of these pigments is expected to be regulated by the transcription factor of flavonoid biosynthetic pathway. Thus, it is important to examine whether *Tamyb10* gene carrying R2R3 MYB domain can regulate the biosynthesis. To investigate the ability of *Tamyb10* to activate flavonoid biosynthetic genes, a transient assay system using white coleoptiles previously developed (Ahmed et al. 2006) was applied. We constructed a plasmid, pBI-myb10D, with the *Tamyb10-D1* gene of CS driven by the CaMV35S promoter and introduced it into white coleoptiles of CS by particle bombardment. Figure 7a shows an obvious red pigmentation in the coleoptiles 24 h after bombardment of *Tamyb10-D1* of CS. The red pigment appeared to be anthocyanin from the absorbance spectra with acidic methanol extract of the pigmented coleoptiles (data not shown). On the other hand, the delivered empty vector (pBI221) did not induce red pigmentation (Fig. 7b). We also examined the expression of flavonoid biosynthetic genes, including *CHS*, *CHI*, *F3H*, *DFR*, and *ANS*, in *Tamyb10-D1*-induced red-pigmented coleoptiles. The expressions of *CHI*, *F3H*, *DFR*, and *ANS* were higher in pigmented coleoptiles than in white coleoptiles, which harbor a control vector (Fig. 7c). It is noteworthy that a particularly high expression of the *DFR* gene was found in pigmented coleoptiles but not in white ones. These results suggest that

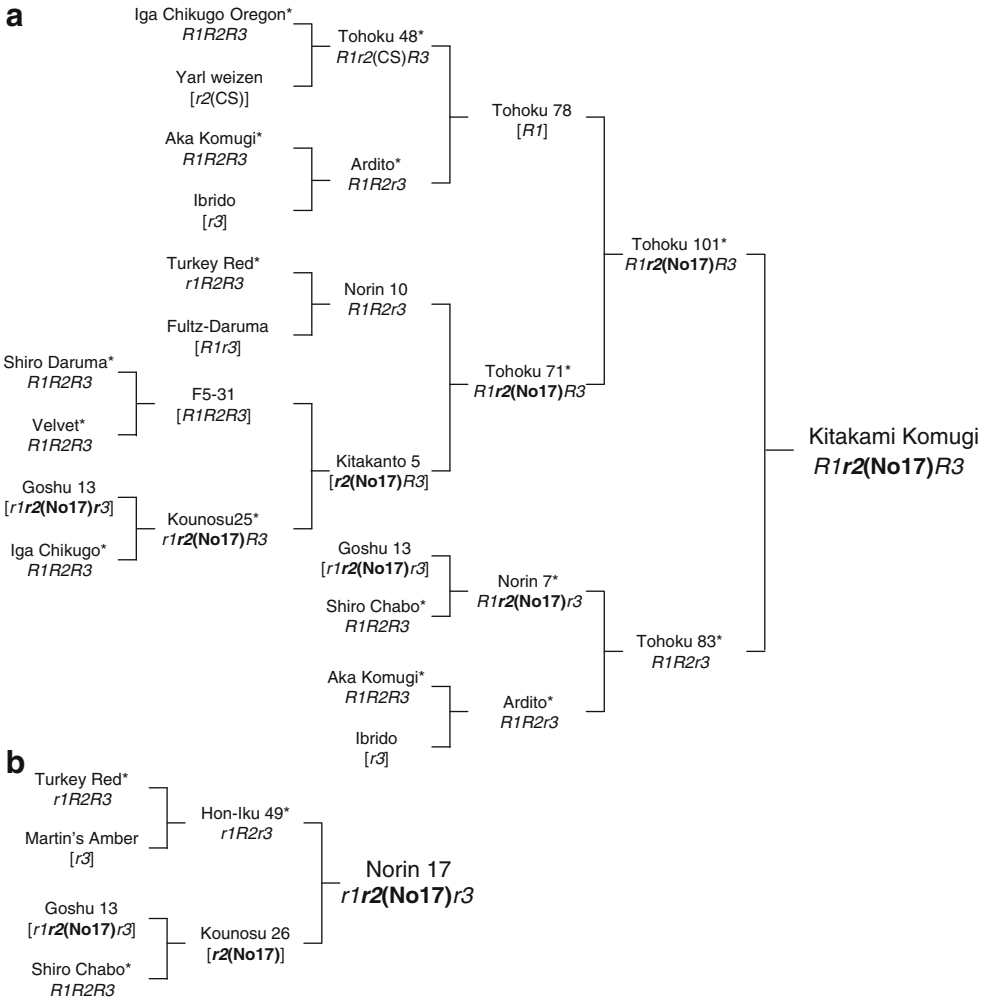

Fig. 6 Pedigree charts of Kitakami Komugi and Norin 17. *R1/r1*, *R2/r2*, and *R3/r3* are synonymous with *R-D1b/R-D1a*, *R-A1b/R-A1a*, and *R-B1b/R-B1a*, respectively. *Tamyb10-A1* genes of CS type are written as "*r2* (CS)", and those of the Norin 17 type are written as "*r2* (No17)" *in bold font*. Estimated *R-1* gene genotypes are shown *within brackets*. The progenitors of Kitakami Komugi (**a**) and Norin 17 (**b**) were cited from Yamada (1990). Wheat varieties with *asterisks* were distributed from a genetic resource of the NIAS (National Institute of Agrobiological Sciences)

the product of *Tamyb10-D1* of CS can induce the expression of flavonoid biosynthetic genes and regulate anthocyanin synthesis.

The function of *Tamyb10-A1* genes of red-grained AUS1490 and its mutant, white-grained EMS-AUS, was also analyzed in our transient system (Fig. 8a). Both *Tamyb10-A1* genes were expressed in immature grain (DPA 5). Whereas *Tamyb10-A1* from AUS1490 made the CS coleoptiles pigmented through a transient assay, that from EMS-AUS did not induce anthocyanin pigmentation (Fig. 8b, c). As reported before, guanine at position 52 of AUS1490 was replaced into adenine in EMS-AUS, resulting in a single amino acid substitution, glycine to glutamic acid, at position 15 in EMS-AUS (Fig. 8d). Glycine at position 15 was extremely conserved among functional MYB proteins, including TT2 (PA regulator) (Nesi et al. 2001), PAP1, PAP2 (anthocyanin regulator) (Borevitz et al. 2000), AtMYB4 (UV-protecting sunscreens regulator) (Jin et al. 2000), AtMYB34 (ATR1, tryptophan pathway regulator) (Bender and Fink 1998), GL1 (AtMYB0, regulator of trichome development) (Oppenheimer et al. 1991), AtMYB66 (WER, regulator of epidermal-cell patterning) (Lee and Schiefelbein 1999), and AtMYB2 (regulator of salt- and dehydration-responsive genes) (Urao et al. 1993) (Fig. 8d). In addition, the glycine site was conserved among 96% of 125 R2R3-MYB proteins of *Arabidopsis* (Fig. 8d) (Stracke et al. 2001). Furthermore, the predicted alpha-helix structure differs between Tamyb10-A1 of AUS1490 and EMS-AUS, as determined using the New Joint method (Akiyama et al. 1998) (data not shown). These findings suggest that the glycine of red-grained AUS1490 is essential as a functional regulator.

Since *Tamyb10-B1* derived from red-grained Norin 61 showed high identity to *Tamyb10-A1* and *Tamyb10-D1*, the product of *Tamyb10-B1* of dominant *R-B1b* alleles presumably has the same active function as that of *Tamyb10-D1* of CS or *Tamyb10-A1* of AUS1490. Consequently, *Tamyb10-A1* and *Tamyb10-D1* derived from red-grained AUS1490 and CS, respectively, were demonstrated to activate anthocyanin biosynthetic enzymes, suggesting that *Tamyb10* is a strong candidate for the *R-1* gene.

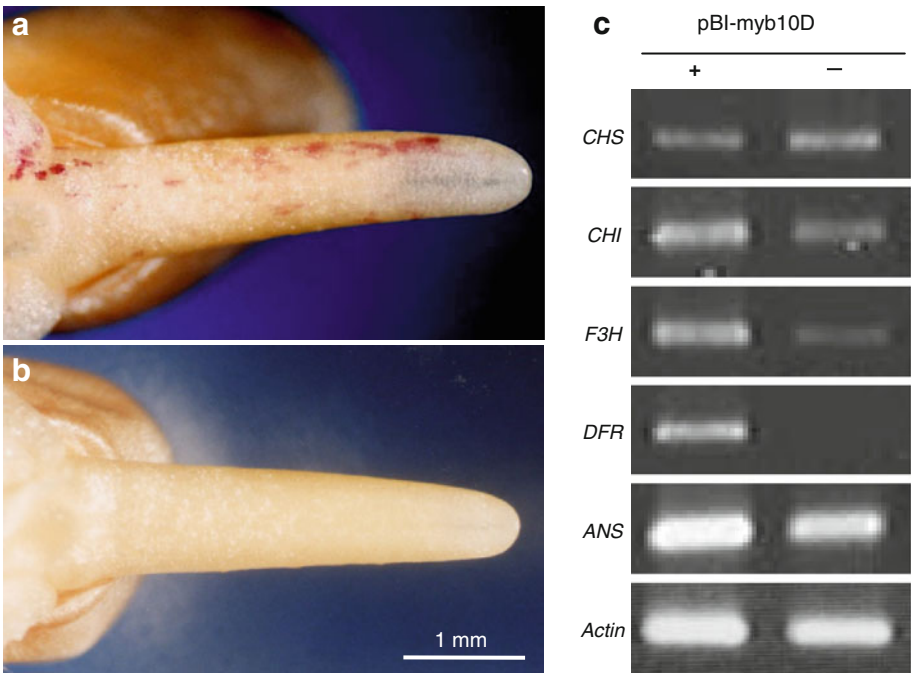

Fig. 7 Pigmentation of the coleoptile by delivering the plasmid, pBI-myb10, that introduced the *Tamyb10-D1* gene of CS under the control of the CaMV 35 promoter. **a** Coleoptile 24 h after bombardment with pBI-myb10. **b** Coleoptile 24 h after bombardment with a plasmid pBI221. *Bar* 1 mm. **c** Expression of flavonoid biosynthetic genes in coleoptiles with *Tamyb10-D1* of CS (+) or pBI-221 (−) bombardment. RNA samples of coleoptiles were isolated 24 h after bombardment and used for RT-PCR

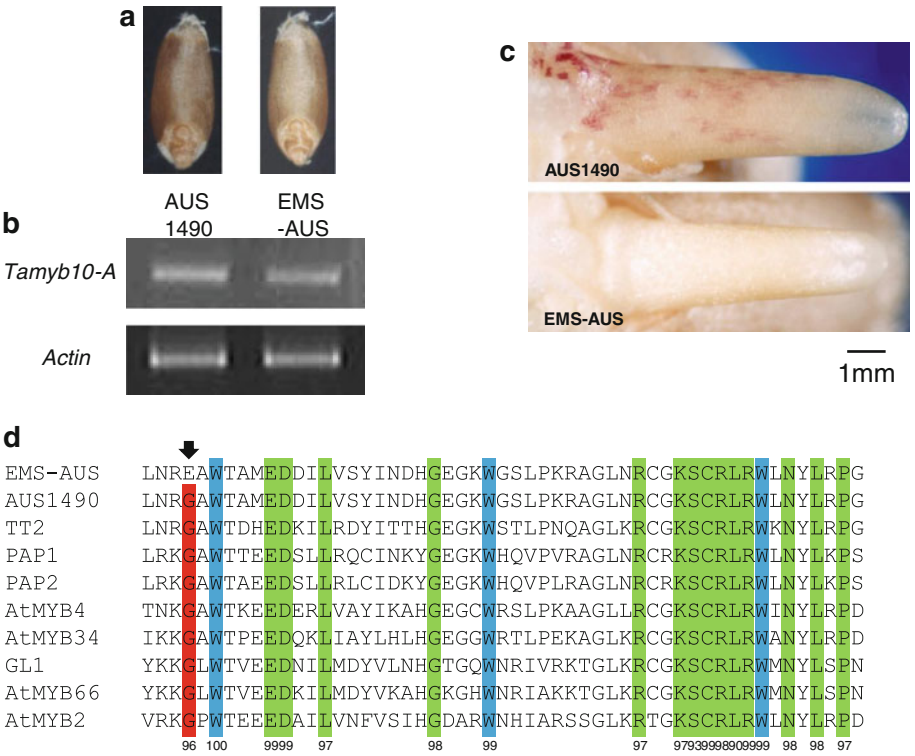

Fig. 8 Functional analysis of Tamyb10-A1 of AUS1490 and EMS-AUS. **a** Grain color of AUS1490 and its mutant, EMS-AUS. **b** Expression of *Tamyb10-A1* in AUS1490 and EMS-AUS. The *actin* gene is used as the control, showing that the same amount of cDNA was applied in the PCR. **c** Pigmentation of the 24-hour coleoptile by delivering the plasmid with the *Tamyb10-A1* gene of AUS1490 (*upper*) or EMS-AUS (*lower*) under control of the CaMV 35 promoter. **d** R2 repeat of MYB protein sequences. The alteration site from conserved glycine (G), colored in *red*, to glutamic acid (E) in EMS-AUS is indicated by a *black arrow*. Repeated W (tryptophan) residues in the R2R3 motif are shown in *blue*. Other conservative residues are shown in *green*. The *numbers at the bottom* are the conservation ratio among 125 R2R3-MYB proteins of *Arabidopsis* (Stracke et al. 2001)

Detection of non-functional genes in anthocyanin biosynthetic pathway through a transient assay of *Tamyb10* in white coleoptile of white-grained varieties

In *Arabidopsis*, mutants impaired in flavonoid accumulation in seeds have been identified as *transparent testa* (*tt*) mutants (Lepiniec et al. 2006). Characterized loci involved in PA biosynthesis in *Arabidopsis* seed have been classified into structural genes (such as *CHS*, *CHI*, *F3H*, *F3′H*, *DFR*, and *ANS*) and regulatory genes (encoding regulatory proteins such as MYB, bHLH, MADS, and WD40) (Lepiniec et al. 2006). Mutants of anthocyanidin-less grain (*ant* mutants) are also known in barley, and one of the ant loci, *ant 18*, was reported as *DFR* gene (Kristiansen and Rohde 1991). However, no mutants of structural genes involved in flavonoid biosynthesis in wheat have been reported. It is still unknown whether mutation(s) occurred in structural gene(s) in white-grained wheat.

We performed a transient assay with pBI-myb10D (CS) into white coleoptiles of 11 white-grained lines to survey malformed structural genes of these lines. As shown in Table 4, red pigmentation was observed in coleoptiles of CS and other 10 white-grained lines, but no pigmentation was found in the coleoptile of Clark's cream by introducing the *Tamyb10* gene alone. On the other hand, delivering both *Tamyb10* and *B-peru*, the bHLH protein of maize induced red pigmentation in the coleoptile (Table 4). These results suggest that all structural genes involved in anthocyanin accumulation in these white-grained lines act normally. However, it was demonstrated that anthocyanin accumulation in coleoptile requires not only a MYB gene, such as *Tamyb10* or *C1*, but also unknown bHLH proteins, such as *B-peru*, which may be inactive in Clark's cream coleoptiles.

Table 4 Anthocyanin pigmentation after bombardment of *Tamyb10-D1* of CS, or *Tamyb10-D1* of CS together with *B-peru* of maize

Varieties/ lines	Grain color	+pBI-myb10D (CS)	+pBI-myb10D (CS) +B-peru
8019R1	White	+	NT
8021V2	White	+	NT
AUS1408	White	+	NT
BL1496	White	+	NT
Cadoux	White	+	+
Clark's cream	White	–	+
Gaines	White	+	+
Hakei 91-64	White	+	+
Norin 17	White	+	NT
Prina	White	+	+
Ryubaku 7	White	+	NT
CS	Red	+	+

+ red pigmentation, − no pigmentation, *NT* not tested

Discussion

The first report of the inheritance of red grain color of hexaploid wheat was published in 1905 (Biffen 1905). Subsequent reports verified the existence of the *R-1* gene for the redness of grain on the chromosome of the homoeologous group 3 and suggested that the *R-1* gene alone controls the grain color, as maize *P* (ZmP) alone regulates red pigment accumulation in the pericarp (Grotewold et al. 1994). ZmP is a member of the R2R3-type MYB transcription factors that control a set of flavonoid biosynthesis genes (*CHS*, *CHI*, and *DFR*) that are independent of bHLH (Grotewold et al. 1994). While we isolated several sequences that have the bHLH motif in wheat, these bHLH genes were expressed in leaves, and no expression of these bHLH genes was found in immature grain (Himi and Noda 2005). Similar results were obtained from the Wheat Genetic Resource Database (http://www.shigen.nig.ac.jp/wheat/komugi/top/top.jsp), suggesting that bHLH proteins are less active than MYB proteins in developing grains.

We isolated the *Tamyb10* genes of wheat, which encode R2R3-type MYB transcription factors. These proteins have a unique motif (IRTKAL/IRC) in the C-terminus region, which was also found in TT2 of *Arabidopsis*, OsMYB3 of rice, and other PA regulators (Fig. 2c; Suppl Fig. S4). The KAxRC sequences in the motif were also seen in the C1 protein, which is also a regulator of the anthocyanin biosynthesis of maize, suggesting that this site might play an essential role as a flavonoid biosynthesis regulator. Although the *Tamyb10-B1* gene was expressed in both *R-B1a* and *R-B1b* varieties (data not shown), the deduced amino acid sequences encoded by *Tamyb10-B1* of CS and other varieties carrying *R-B1a*, a recessive allele, may not have the function because they lack this conserved motif. We demonstrated the critical difference of sequences between two types of *Tamyb10-A1* genes of *R-A1a* varieties and *R-A1b* varieties and the *Tamyb10-B1* genes of *R-B1a* varieties and *R-B1b* varieties. On the other hand, the *Tamyb10-D1* sequence of the *R-D1a* varieties remains to be isolated (Fig. 2a, b). No amplified fragments from genomic DNA and cDNA of *R-D1a* varieties were obtained by PCR using *Tamyb10-D1*-specific primers, suggesting that the *Tamyb10-D1* sequence of *R-D1a* might be deleted or that the sequence of the sites where the primers anneal might be varied. However, the presence/absence of *Tamyb10-D1* is demonstrated through PCR amplification using Tamyb10-D1-specific primers, although heterozygotes of *Tamyb10-D1* cannot be detected by PCR. Consequently, through strong linkage between the *R-1* genotype and the

Tamyb10-sequence type of 33 varieties carrying known *R-1* genotypes and DH plants, Tamyb10-genome-specific primers were found to be efficient for *R-1* genotyping.

Previous reports indicated that the dominant *R-A1b*, *R-B1b*, and *R-D1b* loci are functionally equivalent and have a quantitatively similar effect (Flintham 2000). The putative amino acid sequences of Tamyb10-A1, B1, and D1 of the dominant *R-A1b*, *R-B1b*, and *R-D1b* alleles are similar to each other, and each Tamyb10 appears to act as a transcription factor. On the other hand, *Tamyb10-A1* of CS, which has a recessive *R-A1a* allele, was found to lack the first half of the R2 repeat of the MYB region as if it were bound to another gene through a large deletion, including first half of the R2 repeat (Fig. 2a, b). Interestingly, the promoter sequence of the *Tamyb10-A1* gene isolated from diploid wheat (*Triticum monococcum*, $2n = 14$; genome formula, A^mA^m) showed similarity to that of *Tamyb10-A1* of Norin 61, which has a dominant *R-A1b* allele (data not shown). Furthermore, two mutation patterns of *Tamyb10-A1* of CS type (lack of first half of the R2 repeat) and Norin 17 type (*GeST* insertion in second intron) are not found in one *Tamyb10-A1* gene simultaneously. Presumably, each mutation could result from a different event in the wheat line of the dominant *R-A1b* in an evolutionary step. The 19-bp deletion in the *Tamyb10-B1* gene of the recessive *R-B1a* allele could also result from a single event in the wheat line of the dominant *R-B1b*, as well as *R-A1*.

Recently, more than three thousand transposable elements (TEs) were identified from wheat chromosome 3B (Choulet et al. 2010). While characterized mutable alleles caused by TEs in the *hAT* family have been reported in many species (Kempken and Windhofer 2001), few mutable alleles have been reported in wheat even though the wheat genome contains TEs in the *hAT* family. We indentified a novel transposon, *GeST*, belonging to the *hAT* family, in *Tamyb10-A1* of Norin 17. However, *GeST* has no *hAT* dimerization motif and is locked in the same positions in several varieties or from generation to generation, suggesting that *GeST* is inactive.

Zimmermann et al. (2004) reported a conserved amino acid signature ($D/ELx_2R/Kx_3Lx_6Lx_3R$) as the structural basis for interaction between MYB and bHLH proteins. Interestingly, this signature was found in Tamyb10 proteins (Suppl Fig. S4) but not in ZmP, an independent transcription factor for phlobaphene. In this report, we show that the introduced *Tamyb10* gene induced red anthocyanin pigmentation in white coleoptiles of CS but not in white coleoptiles of Clark's cream (Table 4). We previously showed that the introduction of *C1* (a MYB-type regulatory gene of maize) and *B-peru* (a bHLH-type regulatory gene of maize) induced anthocyanin accumulation in coleoptiles of CS (Ahmed et al. 2003). From these results, not only a MYB gene such as *Tamyb10* or *C1* but also unknown bHLH proteins such as *B-peru*, which may be inactive in Clark's cream coleoptiles, regulate to activate anthocyanin biosynthesis genes in coleoptiles. Khlestkina et al. (2008) reported that *Rc* (red coleoptiles) genes located on homoeologous group 7 are responsible for anthocyanin pigmentation, which regulates *F3H* expression in wheat coleoptiles. The introduction of *Tamyb10* also upregulated *F3H* and other structural genes, suggesting that *Rc* might be a *MYB*-type regulatory gene, such as *Tamyb10*. *Tamyb10* was demonstrated to activate anthocyanin biosynthesis genes with synergy of the bHLH-type protein through a transient assay. However, no bHLH-type proteins have been detected in immature grains of red-grained wheat by bHLH-specific PCR or registered in the wheat database. Furthermore, genetic analysis for red grain has indicated that the genome-specific single gene *R-1* controlled the red pigmentation of grains. This evidence suggests that *Tamyb10* solely activates catechin and PA synthesis genes. From these results, it is likely that *Tamyb10* is a strong candidate for the *R-1* gene.

An association between grain color and depth of dormancy has been noted in a *viviparous 1* (*vp1*) mutant of maize and in *transparent testa* (*tt*) mutants of *Arabidopsis* (McCarty et al. 1991; Debeaujon et al. 2000). The *R-1* genes affect the sensitivity of embryos to abscisic acid (ABA) and the development of grain dormancy, and it has been proposed that one of the *Myb*-type genes of *Arabidopsis*, *AtMYB2*, might be involved in ABA signal transduction (Abe et al. 2003). It remains to be determined whether *Tamyb10* is involved in the development of grain dormancy via control of the sensitivity of the embryo to ABA. The gene structure characterization described in this paper may provide important information related to simple and easy genetic markers for wheat grain color and the association between grain color and grain dormancy.

Acknowledgments We thank Dr. Flintham (John Innes Centre) for providing Novosibirskaya 67 and its NIL; Dr. Mares (The University of Adelaide) for the gift of AUS1490 and EMS-AUS; Dr. Sorrells (Cornell University) for giving Cornell 595; and Dr. Tanio (National Agricultural Research Center, Japan) for providing SZF. This work was supported by grants from the Iijima Memorial Foundation for the Promotion of Food Science and Technology and the Ministry of Agriculture, Forestry, and Fisheries of Japan.

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(1):63–78

Ahmed N, Maekawa M, Utsugi S, Himi E, Ablet H, Rikiishi K, Noda K (2003) Transient expression of anthocyanin in developing

wheat coleoptile by maize *C1* and *B-peru* regulatory genes for anthocyanin synthesis. Breed Sci 53(1):29–34

Ahmed N, Maekawa M, Utsugi S, Rikiishi K, Ahmad A, Noda K (2006) The wheat *Rc* gene for red coleoptile colour codes for a transcriptional activator of late anthocyanin biosynthesis genes. J Cereal Sci 44:54–58

Akagi T, Ikegami A, Tsujimoto T, Kobayashi S, Sato A, Kono A, Yonemori K (2009) DkMyb4 is a myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. Plant Physiol 151(4):2028–2045

Akagi T, Ikegami A, Yonemori K (2010) DkMyb2 wound-induced transcription factor of persimmon (*Diospyros kaki* Thunb.), contributes to proanthocyanidin regulation. Planta 232(5):1045–1059

Akiyama Y, Onizuka K, Nogushi T, Ando M (1998) Parallel protein information analysis (papia) system running on a 64-node pc cluster. In: Proc the 9th genome informatics workshop (GIW'98), 1998. Universal Academy Press, pp 131–140

Allan RE, Vogel OA (1965) Monosomic analysis of red seed color in wheat. Crop Sci 5:474–475

Bender J, Fink GR (1998) A Myb homologue, ATR1, activates tryptophan gene expression in *Arabidopsis*. Proc Natl Acad Sci USA 95(10):5655–5660

Biffen RH (1905) Mendel's laws of inheritance and wheat breeding. J Agric Sci 1:4–48

Bogs J, Jaffe FW, Takos AM, Walker AR, Robinson SP (2007) The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. Plant Physiol 143(3):1347–1361

Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell 12(12):2383–2394

Buck MJ, Atchley WR (2003) Phylogenetic analysis of plant basic helix–loop–helix proteins. J Mol Evol 56(6):742–750

Choulet F, Wicker T, Rustenholz C, Paux E, Salse J, Leroy P, Schlub S, Le Paslier MC, Magdelenat G, Gonthier C, Couloux A, Budak H, Breen J, Pumphrey M, Liu S, Kong X, Jia J, Gut M, Brunel D, Anderson JA, Gill BS, Appels R, Keller B, Feuillet C (2010) Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. Plant Cell 22(6):1686–1701

Debeaujon I, Leon-Kloosterziel KM, Koornneef M (2000) Influence of the testa on seed dormancy, germination, and longevity in arabidopsis. Plant Physiol 122(2):403–414

Flintham JE (2000) Different genetic components control coat-imposed and embryo-imposed dormancy in wheat. Seed Sci Res 10:43–50

Grotewold E, Drummond BJ, Bowen B, 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(3):543–553

Himi E, Noda K (2004) Isolation and location of three homoeologous dihydroflavonol-4-reductase (DFR) genes of wheat and their tissue-dependent expression. J Exp Bot 55(396):365–375

Himi E, Noda K (2005) Red grain colour gene (*R*) of wheat is a myb-type transcription factor. Euphytica 143:239–242

Himi E, Mares DJ, Yanagisawa A, Noda K (2002) Effect of grain colour gene (*R*) on grain dormancy and sensitivity of the embryo to abscisic acid (ABA) in wheat. J Exp Bot 53(374):1569–1574

Himi E, Nisar A, Noda K (2005) Colour genes (*R* and *Rc*) for grain and coleoptile upregulate flavonoid biosynthesis genes in wheat. Genome 48(4):747–754

Huang J, Zhang K, Shen Y, Huang Z, Li M, Tang D, Gu M, Cheng Z (2009) Identification of a high frequency transposon induced by tissue culture, *nDaiZ*, a member of the *hAT* family in rice. Genomics 93(3):274–281

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(22):6150–6161

Kaneko S, Komae K, Nagamine T, Yamada T (1994) Development of an extraction procedure for wheat red coat pigments and determination of varietal differences for this trait. Breed Sci 44:263–266

Kempken F, Windhofer F (2001) The *hAT* family: a versatile transposon group common to plants, fungi, animals, and man. Chromosoma 110(1):1–9

Khlestkina EK, Roder MS, Salina EA (2008) Relationship between homoeologous regulatory and structural genes in allopolyploid genome—a case study in bread wheat. BMC Plant Biol 8:88

Koval SF (1997) The catalog of near-isogenic lines of novosibirskaya-67 common wheat and principles of their use in experiments. Russ J Genetics 33(8):995–1000

Kristiansen KN, Rohde W (1991) Structure of the *Hordeum vulgare* gene encoding dihydroflavonol-4-reductase and molecular analysis of *ant*18 mutants blocked in flavonoid synthesis. Mol Gen Genet 230(1–2):49–59

Lee MM, Schiefelbein J (1999) WEREWOLF, a MYB-related protein in *Arabidopsis*, is a position-dependent regulator of epidermal cell patterning. Cell 99(5):473–483

Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol 57:405–430

Mares D, Mrva K, Cheong J, Williams K, Watson B, Storlie E, Sutherland M, Zou Y (2005) A QTL located on chromosome 4A associated with dormancy in white- and red-grained wheats of diverse origin. Theor Appl Genet 111(7):1357–1364

Mares D, Rathjen J, Mrva K, Cheong J (2009) Genetic and environmental control of dormancy in white-grained wheat (*Triticum aestivum* L.). Euphytica 168:311–318

Matus-Cadiz MA, Daskalchuk TE, Verma B, Puttick D, Chibbar RN, Gray GR, Perron CE, Tyler RT, Hucl P (2008) Phenolic compounds contribute to dark bran pigmentation in hard white wheat. J Agric Food Chem 56(5):1644–1653

McCallum JA, Walker JRL (1990) Proanthocyanidins in wheat bran. Cereal Chem 67(3):282–285

McCarty DR, Hattori T, Carson CB, Vasil V, Lazar M, Vasil IK (1991) The *Viviparous-1* developmental gene of maize encodes a novel transcriptional activator. Cell 66:895–905

McIntosh RA, Hart GE, Devos KM, Gale MD, Rogers WJ (1998) Catalogue of gene symbols for wheat. In: Proceedings of the ninth international wheat genetics symposium, Canada, 1998. University of Saskatchewan Extension Press

Mellway RD, Tran LT, Prouse MB, Campbell MM, Constabel CP (2009) The wound-, pathogen-, and ultraviolet B-responsive *MYB134* gene encodes an R2R3 MYB transcription factor that regulates proanthocyanidin synthesis in poplar. Plant Physiol 150(2):924–941

Metzger RJ, Silbaugh BA (1970) Location of genes for seed coat color in hexaploid wheat, *Triticum aestivum* L. Crop Sci 10:495–496

Miura H, Sato N, Kato K, Amano Y (2002) Detection of chromosomes carrying genes for seed dormancy of wheat using the backcross reciprocal monosomic method. Plant Breed 121:394–399

Miyamoto T, Everson EH (1958) Biochemical and physiological studies of wheat seed pigmentation. Agron J 50:733–734

Mol J, Grotewold E, Koes R (1998) How genes paint flowers and seeds. Trends Plant Sci 3(6):212–217

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8(19):4321–4325

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(9):2099–2114

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(3):483–493

Quinde-Axtell Z, Baik BK (2006) Phenolic compounds of barley grain and their implication in food product discoloration. J Agric Food Chem 54(26):9978–9984

Ramsay NA, Glover BJ (2005) MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci 10(2):63–70

Routaboul JM, Kerhoas L, Debeaujon I, Pourcel L, Caboche M, Einhorn J, Lepiniec L (2006) Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. Planta 224(1):96–107

Sears ER (1944) Cytogenetic studies with polyploid species of wheat. II. Additional chromosomal aberrations in *Triticum vulgare*. Genetics 29:232–246

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

Styles ED, Ceska O (1975) Genetic control of 3-hydroxy- and 3-deoxy-flavonoid in *Zea mays*. Phytochemistry 14:413–415

Suzuki A, Suzuki T, Tanabe F, Toki S, Washida H, Wu CY, Takaiwa F (1997) Cloning and expression of five *myb*-related genes from rice seed. Gene 198(1–2):393–398

Terrier N, Torregrosa L, Ageorges A, Vialet S, Verries C, Cheynier V, Romieu C (2009) Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. Plant Physiol 149(2):1028–1041

Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K (1993) An Arabidopsis myb homolog is induced by dehydration stress, and its gene product binds to the conserved MYB recognition sequence. Plant Cell 5(11):1529–1539

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC (1999) The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. Plant Cell 11(7):1337–1350

Warner RL, Kudrna DA, Spaeth SC, Jones SS (2000) Dormancy in white-grain mutants of Chinese Spring wheat (*Triticum aestivum* L.). Seed Sci Res 10:51–60

Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol 126(2):485–493

Xie DY, Jackson LA, Cooper JD, Ferreira D, Paiva NL (2004) Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol-4-reductase from *Medicago truncatula*. Plant Physiol 134(3):979–994

Yamada T (1990) Classification of GA response, *Rht* genes and culm length in Japanese varieties and landraces of wheat. Euphytica 50:221–239

Yoshida K, Iwasaka R, Kaneko T, Sato S, Tabata S, Sakuta M (2008) Functional differentiation of *Lotus japonicus* TT2s, R2R3-MYB transcription factors comprising a multigene family. Plant Cell Physiol 49(2):157–169

Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF (2004) Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like bHLH proteins. Plant J 40(1):22–34