CHAPTER III

IDENTIFICATION OF ETHANOL-INDUCIBLE GENES AND ISOLATION OF THE *MYB-RELATED PROTEIN-LIKE* PROMOTER IN *ORYZA SATIVA* L.

3.1 Introduction

Rice (*Oryza sativa* L.) is a rich source of phytonutrients and antioxidants. Rice has been cultivated and consumed mainly in Asia (Lim 2012). Many rice cultivars have evolved stress-tolerant traits to survive under various stress conditions causing lower yield and grain quality. Stress-tolerant traits of rice such as drought, disease, and cold tolerance are controlled by genes involved in stress response systems (Pérez-Torres et al. 2009). Nutritions available in rice grain are also products of gene functions in rice. In addition, special characteristics of rice including antioxidant compounds and aroma are also regulated by genes involved in different pathways (Vanavichit and Yoshihashi 2010). Understanding regulation of these genes is necessary for crop improvement. Moreover, rice is an important model plant for functional genomic analysis of monocots, especially cereals (Ohnishi et al. 2011). The genomes of cereals including rice, wheat, sorghum, maize, barley, and millet are highly conserved (Eckardt 2001). Therefore, knowledge of gene function in rice could be widely applied for crop improvement in other cereals.

The regulatory region or 'promoter' that is used to control target genes plays a key role in gene expression studies (Komarnytsky and Borisjuk 2003). Several promoters including constitutive, tissues-specific, and chemical-inducible promoters are generally used to regulate genes of interest in model plants (Li et al. 2005). Constitutive promoters such as the cauliflower mosaic virus 35S (CaMV 35S), ubiquitin, and actin promoters drive a constitutive expression of target genes in all

tissues and developmental stages of plants (Gurr and Rushton 2005). However, overexpression of the target genes can lead to deleterious effects on plant morphology and health (Fu et al. 2001). To overcome this problem, tissue-specific promoters could be used. Tissue-specific promoters confine gene expression to specific tissues or specific developmental stage of plants (Gurr and Rushton 2005), but the limitations are that only few tissue-specific promoters are available, and they are restricted to only a few cell types in some plants (Li et al. 2005). To resolve the limitations of constitutive and tissue-specific promoters, many chemical-inducible systems have been developed for gene expression analysis.

Chemical-inducible systems consist of two transcription units (Padidam 2003). The first unit encodes a transcription factor that responds to chemical inducers. Then, an activated transcription factor only binds to its response element in the second transcription unit and activates gene transcription. Chemicals that are used to induce transcription of target genes include dexamethasone (Aoyama and Chua 1997), estrogen (Zuo et al. 2000), tetracycline (Weinmann et al. 1994), ecdysone (Martinez et al. 1999), copper (Mett et al. 1993), and ethanol (Caddick et al. 1998). Among them, ethanol is an ideal inducer since it is biodegradable, cheap, and environmentally safe. Furthermore, only small amount of ethanol is required to activate the ethanol-inducible system in plants in which visible phytotoxic symptoms are not observed (Ait-ali et al. 2003; Caddick et al. 1998).

An ethanol-inducible system or '*alc* switch' comprises a constitutively expressed *alcR* gene that encodes an alcohol-regulated transcription factor (ALCR) and the ethanol-inducible promoter isolated from the *alcA* gene that encodes alcohol dehydrogenase I. ALCR is in an inactive form in the absence of ethanol.

Under the ethanol treatment, ALCR is activated. The activated-ALCR binds to a response element in *alcA* promoter and activates gene transcription (Felenbok et al. 1988). The *alc* switch has been used to regulate several genes in various plant species including a semi-dwarfing gene in transgenic Arabidopsis (Ait-ali et al. 2003) and chloramphenicol acetyltransferase (CAT)-encoding gene in tomato (Garoosi et al. 2005) and *Catharanthus roseus* (Peebles et al. 2007). Moreover, the ethanol-induced expression of *TBP-associated factor* (*TAF*) could restore the male fertility in the male sterile transgenic eggplant (Toppino et al. 2011). This information indicates the feasibility of an ethanol-inducible system to spatially and temporally regulate a wide range of genes. However, this system is derived from fungus, *Aspergillus nidulans* which has a substantial genetic distance from plants (Felenbok et al. 1988), and two transcription units of the *alc* switch must be transformed into the transgenic plants (Padidam 2003). Therefore, the use of a single transformation unit containing an ethanol-inducible promoter derived from plants leading to regulation of a target gene expression is of interest.

Recently, the ethanol responses of non-transformed plants have been discovered. Vreugdenhil *et al.* (2006) reported that ethanol broke a dormancy period of non-transformed potato tubers and stimulated sprouting. Their study on gene expression analysis revealed that ethanol down-regulated the expression of genes involved in cell division and storing reserves in ethanol-treated potatoes (Vreugdenhil et al. 2006). Sugarcane is another example of a non-transformed plant that responds to ethanol. The transcription profiles of genes in sugarcane leaves were shown to be altered in the presence of ethanol, and the ethanol-responsive genes were identified (Camargo et al. 2007). The responses of normal plants to exogenously applied

ethanol suggested that higher plants including rice may have all the components needed for ethanol regulation. Thus, it is possible to use just an ethanol-inducible promoter from a rice plant to construct a single transformation unit to regulate expression of a desirable gene. This may function as well as the *alcR-alcA* switch from *A. nidulans*, but much easier for generation of desirable transgenic plants.

Therefore, this study was aimed to identify ethanol-inducible genes in young panicles of rice plants treated with ethanol using cDNA-AFLP. The expression of ethanol-inducible genes was confirmed by semi-quantitative reverse transcription-polymerase chain reaction (semi-quantitative RT-PCR). Furthermore, the promoter region or 5' upstream region (UTR) was isolated from a highly conserved ethanol-inducible gene. The *cis*-acting regulatory elements were identified in the isolated 5' UTR using databases of plant *cis*-acting regulatory elements, PLACE and PlantCARE. The information obtained from the present study provides an alternative way to control gene expression. Moreover, this knowledge could be useful for generation of transgenic cereal crop plants with desirable traits.

3.2 Materials and methods

3.2.1 Chemicals and reagents

The chemicals and reagents used in all experiments were of molecular and analytical grade. Absolute ethanol was purchased from Merck Millipore (Merck, Darmstadt, Germany). The Magnetic mRNA Isolation Kit (Cat. no. NEB#S1550S) and T4 DNA ligase (Cat. no. M0202S) were obtained from New England Biolabs (New England Biolabs, Ipswich, MA, USA). The RevertAid™ M·MuLV reverse transcriptase, *Eco*RI, *Mse*I, GeneRuler DNA Ladder Mix, and GeneRuler 100 bp and 100 bp plus DNA were

purchased from Thermo Sciencetific (Thermo Sciencetific, Wilmington, DE, USA). The NucleoSpin® Gel and PCR Clean-up Kit (Cat. no. 740609.50) was purchased from MACHEREY-NAGEL (MACHEREY-NAGEL, Düren, Germany). The TRI Reagent® was purchased from Molecular Research Center (Molecular Research Center, OH, USA). The pCR™2.1-TOPO® vector (Cat. no. 450641) was obtained from Invitrogen (Invitrogen, Carlsbad, CA, USA). The High-Speed Plasmid Mini Kit (Cat. no. PD100) was purchased from Geneaid (Geneaid, New Taipei, Taiwan).

3.2.2 Plant materials and growth conditions

Oryza sativa L. ssp. *Indica* cultivar 'Pathumthani 1' (PTT1) was used in this study. Seeds were obtained from the Bureau of Seed Multiplication of the Rice Department of Thailand in Bangkok. The seeds were planted in plastic pots filled with 3 kg soil and the plants were grown in natural greenhouse at National Center for Genetic Engineering and Biotechnology (BIOTEC), Thailand Science Park, Pathum Thani, Thailand.

3.2.3 Ethanol treatment and sample collection

For effects of ethanol on rice growth at the seedling stage, twenty-six-day-old rice seedlings were sprayed with tap water (control), 1, 2, 4, 6, or 8% ethanol until solutions dripped off once a day for 7 days. During the treatment, seedlings were harvested after spraying for 2, 5, and 7 days. Shoot dry weight and shoot length were measured. Three independent experiments were performed. One hundred and eighty plants were used for each experiment (n = 10 per treatment).

For effects of ethanol on rice growth at maturing, one-hundred and sixteen-day-old rice plants at the vegetative (tillering) stage prior panicle initiation were sprayed with tap water, 1, or 2% ethanol until solutions dripped off once a day for 7 days. In each treatment, rice plants were divided into three groups. In the first group, one day later after the treatment, young panicles were individually collected from five rice plants. Then one panicle from each plant was pooled together, frozen immediately in liquid nitrogen, and stored at -80°C for further study on gene expression analysis. For the second group, one day later after the treatment, rice plants were harvested, and agronomic traits such as number of tillers, shoot height, shoot dry weight, number of panicles, and panicle length were measured. And the last group, rice plants were grown under normal condition after the treatment and were harvested at the ripening stage. Agronomic traits of these plants including the number of tillers, shoot height, shoot dry weight, days to flowering, seed setting rate, and grain yield were determined. Three independent experiments were performed. Fifteen plants were used for each experiment (n = 5 per treatment).

For tissue specific expression, all tissues were harvested from PTT1 rice plants grown under normal condition including, shoots from 15-day-old seedlings, roots, stems, leaf sheaths, leaf blades, young panicles and anthers from rice plants at the vegetative stage, each from five plants. Calli of PTT1 were generated on N6D medium under a sterile condition. Calli were grown under fluorescent light (16/8 photoperiod) at 32°C and were harvested after the third subculture. Samples were frozen immediately in liquid nitrogen and stored at -80°C.

3.2.4 cDNA-AFLP analysis

mRNA was isolated from young panicles of control and 1% ethanol-treated rice plants using the Magnetic mRNA Isolation Kit. Immobilized cDNA was synthesized using RevertAid^{TM} M-MuLV reverse transcriptase. Oligo $d(T)_{25}$ magnetic beads from the Magnetic mRNA Isolation Kit were used as primers for cDNA synthesis. cDNA-AFLP analysis was performed according to the method described by Bachem et al. with minor modifications (Bachem et al. 1998). Immobilized cDNA was digested with 10 units of *Mse*I at 65°C for 1 h. Then immobilized cDNA with *Mse*I-cut site at the 3' end was digested with 10 units of *Eco*RI at 37°C for 1 h. Doubly-digested cDNA was separated from singly-digested immobilized cDNA and ligated to 3 pmoles of *Eco*RI adaptor and 30 pmoles of *Mse*I adaptor (**Table 3-1**) using T4 DNA ligase. The pre-selective amplification was carried out using pre-selective primers corresponding to *Eco*RI and *Mse*I adaptors (**Table 3-1**). Thermal conditions for pre-selective amplification were as follows: 20 cycles of 94°C for 30 s, 56°C for 1 min, and 72°C for 1 min. The pre-selective amplification product was diluted 50-fold. Then, an aliquot of 2 µl was used for selective amplification. A total of 214 combinations of *Eco*RI and *Mse*I selective primers, with selective nucleotides (NN) added to the 3' end, were used (**Appendix, Table A1**). Touchdown program (12 cycles of 94°C for 30 s, 65°C for 30 s which decreased 0.7°C every cycle, and 70°C for 1 min) followed by 23 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 1 min was used for the selective amplification. The selective amplification products were separated using 6% denaturing polyacrylamide gels. Transcript-derived fragments (TDFs) were visualized by silver staining.

Table 3-1 Sequences of adaptors and pre-selective primers used in this study

Adaptors and primers	Sequences (5'⟶3')
*Eco*RI top adaptor	CTCGTAGACTGCGTACC
*Eco*RI bottom adaptor	AATTGGTACGCAGTC
*Mse*I top adaptor	GACGATGAGTCCTGAG
*Mse*I bottom adaptor	TACTCAGGACTCAT
*Eco*RI pre-selective primer (E0)	GACTGCGTACCAATTC
*Mse*I pre-selective primer (M0)	GATGAGTCCTGAGTAA

3.2.5 Isolation of TDFs, DNA sequencing, and BLAST search

The up-regulated and down-regulated TDFs were cut from polyacrylamide gel and eluted in distilled water at 50°C for 1 h. The re-amplification of polymorphic TDFs was performed using the same selective primer combination. Thermal conditions were as follows: 94°C for 4 min, 30 cycles of 94°C for 30 s, 56°C for 1 min, 72°C for 1 min, and 72°C for 7 min. The PCR products were separated in 1% agarose gels, and were purified using the NucleoSpin® Gel and PCR Clean-up Kit. The purified products were sent to First BASE Laboratories (Selangor, Malaysia) for nucleotide sequencing. The resulting sequences were analyzed for homology against Gramene database (http://www.gramene.org/) using BLAST search.

3.2.6 Expression analysis by semi-quantitative RT-PCR

For confirmation of some selected differential TDFs obtained from cDNA-AFLP, total RNA was isolated from young panicles of control and 1% ethanol-treated rice plants. For expression in different tissues of *Myb-related protein-like* (*Os07g0627300*), total RNA was extracted from callus, rice seedlings, and several tissues including roots, stems, leaf sheaths, leaf blades, young panicles, and anther of rice plants grown under normal condition. Total RNA was isolated using TRI Reagent® according to the method described by Chomczynski (Chomczynski 1993). Oligo $d(T)_{18}$ was used as a primer in cDNA synthesis. Semi-quantitative RT-PCR was performed using gene specific primers corresponding to the TDFs obtained from cDNA-AFLP. The sequences of specific primers, annealing temperature and number of PCR cycles used for the selected genes are listed in **Table A2 (Appendix)**. The PCR products were separated in 1% agarose gel and were visualized with ethidium bromide. Three independent experiments were performed for each gene. The amplification of rice *act1* gene was used as an internal control.

3.2.7 Phylogenetic analysis

Phylogenetic analysis was performed using deduced amino acid sequences of Myb-related proteins from rice and other similar sequences from different species across the kingdom obtained from NCBI Reference Sequence database (http://www.ncbi.nlm.nih.gov/refseq/). The tree was generated by the neighbor-joining (NJ) algorithm with p-distance and pairwise deletion of gap, employing in MEGA version 5 program (Tamura et al. 2011). A bootstrap statistical analysis was performed with 1000 replicates to test the phylogeny.

3.2.8 Amplification and sequencing of the 5' upstream region of *Myb-related protein-like* (*Os07g0627300*)

DNA was extracted from leaves of PTT1 rice plants grown under normal condition using the cetyltrimethylammonium bromide method as described previously (Kistler 2012). About 2,000 bp-DNA sequences of the 5' upstream region (UTR) of *Os07g0627300* gene in *Japonica* rice (nipponbare) and *Indica* rice obtained from Gramene were aligned using Clustal W version 1.4 (Thompson et al. 1994). Specific primers were designed based on the conserved sequence of *Indica* and *Japonica* rice. The 5' UTR was amplified using specific primers P627300F1 and P627300R1 (**Table 3-2**). Thermal conditions were as follows: 94°C for 4 min, 35 cycles of 94°C for 45 sec, 69°C for 45 sec, 72°C for 2.30 min, and 72°C for 10 min. The PCR products were separated in a 1% agarose gel and were visualized with ethidium bromide. The DNA fragment was sliced from agarose gel and purified using the NucleoSpin® Gel and PCR Clean-up Kit. The 5' UTR was cloned into pCR™2.1-TOPO® vector, and transformed using *E. coli* DH5α. Positive colonies from selectable media were screened by colony PCR using specific primers to the 5' UTR sequence, P627300F2 and P627300R2 (**Table 3-2**). Fifteen positive clones from the PCR were grown in LB-ampicillin broth and the plasmids were prepared using the High-Speed Plasmid Mini Kit. Plasmids were checked again by PCR using specific primers P627300F3 and P627300R1-2 (**Table 3-2**) which cover the middle region of the fragment. Seven positive plasmids were sequenced using specific primers to the 5' UTR sequence and pCR™2.1-TOPO® vector (**Table 3-2**).

Table 3-2 Primers used in the amplification and sequencing of the 5' upstream region of *Os07g0627300* gene from genomic DNA of PTT1 cultivar

Primers	Sequences (5'⟶3')
Forward strand	
P627300F1	ACGGCGGTGAGCTTACCAAT
P627300F2	CAAACTTGAAACAATTTGATTTTGAC
P627300F3	GCATTGCCGAACAGACCCAT
M13 Reverse	CAGGAAACAGCTATGAC
Reverse strand	
P627300R1	GGTCCGGGATCCCGCATGTT
P627300R1-2	TCATACCGGTGGTCGAATAATGA
P627300R2	CGACTCCCTACAGAATGGCG
M13 Forward (-20)	GTAAAACGACGGCCAG

3.2.9 Sequence analysis

The *cis*-acting regulatory elements in the 5' UTR sequence isolated from PTT1 were predicted using PLACE (https://sogo.dna.affrc.go.jp/cgi-bin/sogo.cgi) and PlantCARE databases (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

3.2.10 Statistical Analysis

Results are reported as means ± SE. The data were analyzed using Tukey's honestly significant difference (HSD) test in one-way analysis of variance (ANOVA) using SPSS software, version 16.0 (SPSS Inc., Chicago, IL). Significant differences were considered at $p < 0.05$.

3.3 Results

3.3.1 Effects of ethanol on growth of rice plants

To determine effects of exogenously applied ethanol on rice growth, rice plants at seedling and reproductive stage were subjected to ethanol treatment. Twenty-six-day-old seedlings were exceedingly sprayed with tap water (control), 1, 2, 4, 6, or 8% ethanol once a day for 7 days. During the treatment, seedlings were harvested after spraying for 2, 5, and 7 days. Shoot dry weight and shoot length of the seedlings were measured. The experiments were performed for three times. The results showed that, 2 days after ethanol application, all concentrations of ethanol had no effects on shoot dry weight (**Figure 3-1a**) and shoot length (**Figure 3-1b**). The effects of ethanol on growth of rice seedlings were observed after 5 and 7 days of ethanol application. Shoot dry weight and shoot length were significantly decreased after application of ethanol higher than 4% ($p < 0.05$) for 5 days. After 7 days of application, ethanol showed stronger effects; applications of ethanol with 2% and above showed the inhibitory effects on growth of rice seedlings both for shoot dry weight and shoot length (**Figure 3-1**). Therefore, ethanol at the concentration of 1 and 2% were used to determine their effects on growth of rice plants at the reproductive (the panicle initiation) stage.

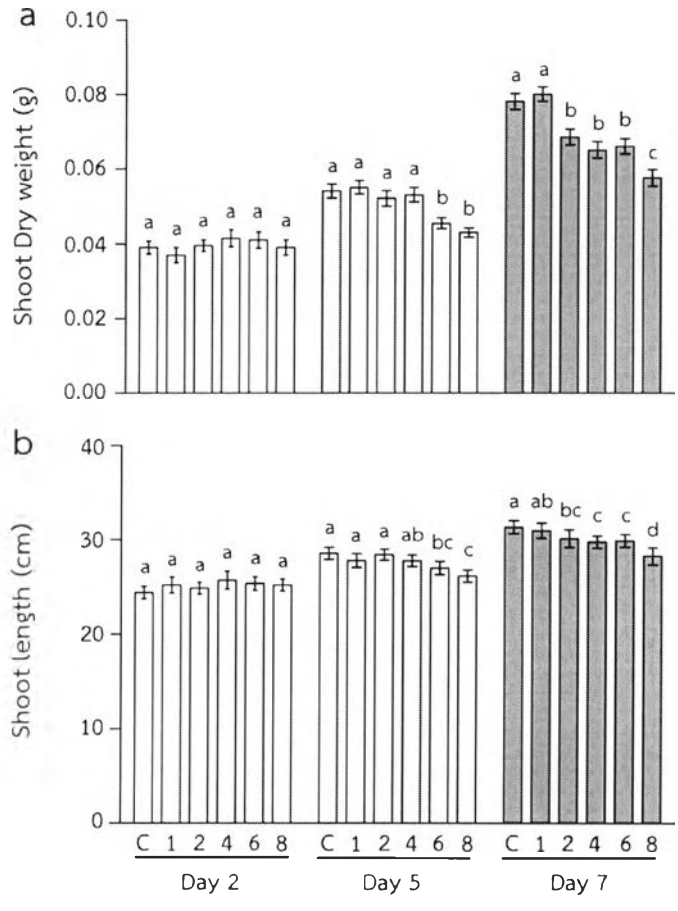

Figure 3-1 Effects of ethanol on shoot dry weight and shoot length of rice at the seedling stage. Twenty-six-day-old seedlings were sprayed daily with tap water (control), 1, 2, 4, 6, or 8% ethanol for 7 days. Seedlings were harvested at 2, 5, and 7 days during the treatments. Shoot dry weight (**a**) and shoot length (**b**) were measured. The results shown are representative of three independent experiments. Error bars represent SE (n = 10). Columns with the same letters in each experiment are not significantly different at $p < 0.05$. C: control and 1: 1%, 2: 2%, 4: 4%, 6: 6%, and 8: 8% ethanol.

One-hundred and sixteen-day-old rice plants at the vegetative (tillering) stage prior to panicle initiation were exceedingly sprayed with tap water, 1 or 2% ethanol once a day for 7 days. The rice plants were harvested and their agronomic traits were measured. The experiments were performed for three times. The results showed that 1% ethanol did not affect agronomic traits of rice plants (**Figure 3-2**). However, 2% ethanol significantly retarded growth of rice plants resulting in reduction in the number of tillers (**Figure 3-2a**), shoot height (**Figure 3-2b**), shoot dry weight (**Figure 3-2c**), number of panicles (**Figure 3-2d**), and panicle length (**Figure 3-2e**).

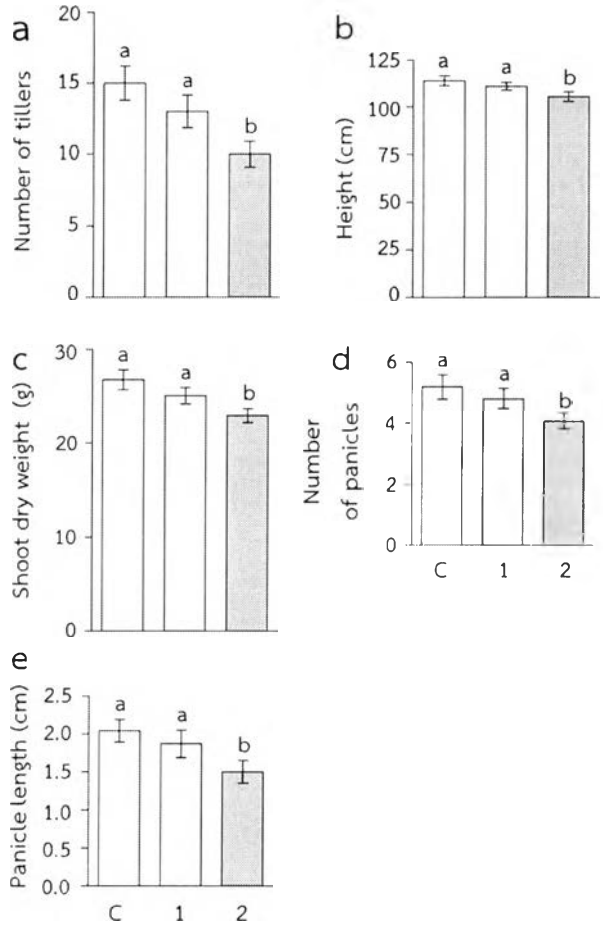

Figure 3-2 Effects of ethanol on growth of rice plants at the reproductive stage (panicle initiation). One-hundred and sixteen-day-old rice plants at the vegetative (tillering) stage prior to panicle initiation were sprayed daily with tap water (control), 1, or 2% ethanol for 7 days. The rice plants were harvested one day later after the treatment. Agronomic traits such as the number of tillers, shoot height, shoot dry weight, number of panicles, and panicle length were measured. (**a**) The number of tillers, (**b**) shoot height, and (**c**) shoot dry weight, and (**d**) number and (**e**) length of panicles of rice plants. The results shown are representative of three independent experiments. Error bars represent SE (n = 5). Columns with the same letters in each experiment are not significantly different at $p < 0.05$. C: control and 1: 1% and 2: 2% ethanol.

Furthermore, the effects of ethanol on yield and other agronomic traits of rice plants at the ripening stage were also examined. One-hundred and sixteen-day-old rice plants at the tillering stage prior to panicle initiation were exceedingly sprayed with tap water, 1 or 2% ethanol once a day for 7 days. After the treatment, rice plants were grown under normal condition and harvested at the ripening stage. Agronomic traits such as the number of tillers, shoot height, shoot dry weight, days to flowering, seed setting rate, and grain yield were determined. 1% ethanol showed no effects on the grain yield, seed setting rate, and other agronomic traits (**Figure 3-3**). However, 2% ethanol affected the number of tillers (**Figure 3-3a**), shoot height (**Figure 3-3b**), shoot dry weight (**Figure 3-3c**), days to flowering (**Figure 3-3d**), seed setting rate (**Figure 3-3e**), and grain yield (**Figure 3-3f**) of rice plants. Based on these results, 1% ethanol was used for further study on effects of ethanol on gene expression.

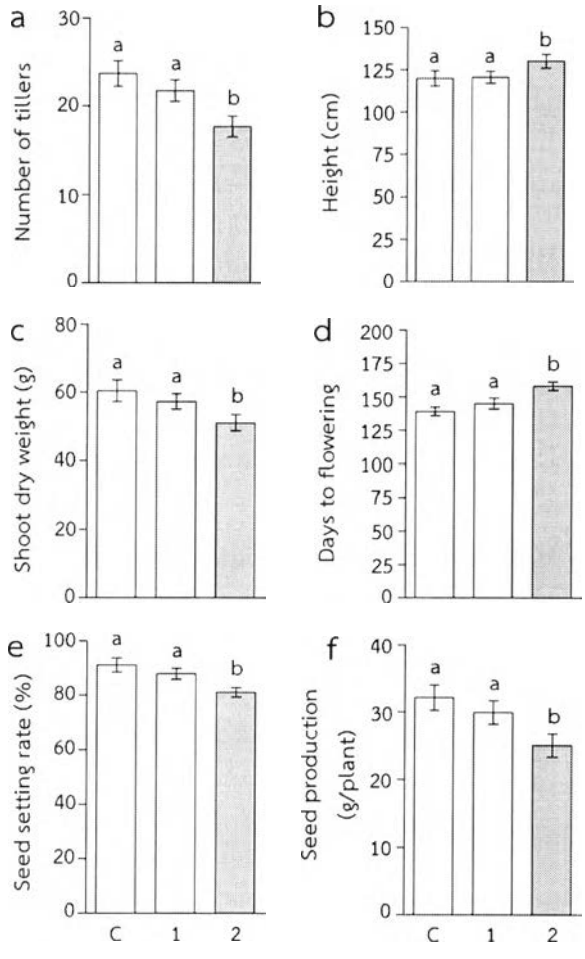

Figure 3-3 Effects of ethanol on yield and other agronomic traits of rice plants at the ripening stage. One-hundred and sixteen-day-old rice plants at the tillering stage prior to panicle initiation were sprayed daily with tap water (control) or 1 and 2% ethanol for 7 days. After the treatment, rice plants were grown under normal condition and harvested at the ripening stage. Agronomic traits such as (**a**) number of tillers, (**b**) shoot height, (**c**) shoot dry weight, (**d**) days to flowering, (**e**) seed setting rate (%), and (**f**) seed production (g/plant) of rice plants were determined. The results shown are representative of three independent experiments. Error bars represent SE. Columns with the same letters in each experiment are not significantly different at $p< 0.05$. C: control and 1: 1% and 2: 2% ethanol.

3.3.2 Identification of ethanol-inducible genes

In order to identify ethanol-inducible genes in young rice panicles, cDNA-AFLP analysis was performed. The rice plants were exceedingly sprayed with tap water (control) or 1% ethanol once a day for 7 days. The young rice panicles of 10 - 15 mm were collected in the next day after the treatment and were used for cDNA-AFLP analysis. The transcript profile of the 1% ethanol-treated young panicle was compared with that treated with tap water, used as control. Samples of cDNA-AFLP profiles derived from *Eco*RI and *Mse*I selective primers is shown in **Figure 3-4**. Two-hundred and eighty-nine of the up-regulated and down-regulated TDFs were selected based on their presence or absence in the cDNA-AFLP profiles. All the TDFs were eluted from polyacrylamide gels and were re-amplified using the same selective primer combinations. After the re-amplification, 99 fragments having the sizes between 400 – 500 bp were chosen for sequencing. Only 34 of these TDFs were successfully sequenced. Sequence homology of the obtained differential TDFs was analyzed using BLAST search in the Gramene rice database (http://www.gramene.org). Twenty-three of the 34 TDFs were found to be homologous to genes of known functions involved in metabolism, biosynthesis, signal transduction, transcriptional regulation, stress response, membrane transport, and cell cycle. Meanwhile, 11 TDFs showed sequence homology with genes of uncharacterized functions (**Table 3-3**). Among the 34 TDFs, 29 were up-regulated while 5 were down-regulated.

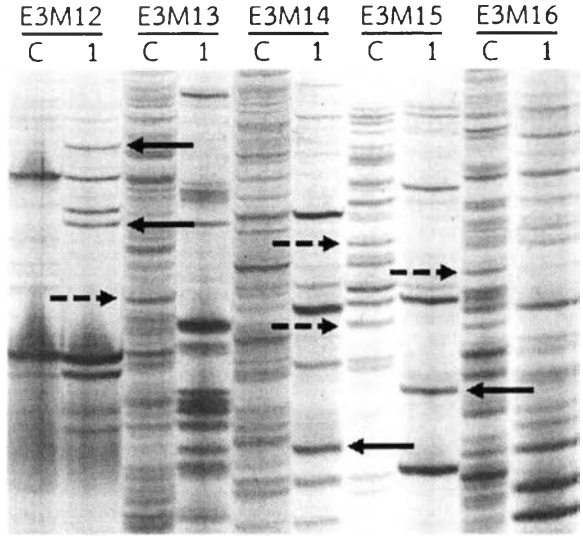

Figure 3-4 Samples of cDNA-AFLP profiles showing differential transcript-derived fragments (TDFs). One-hundred and sixteen-day-old rice plants at the tillering stage prior to panicle initiation were sprayed daily with tap water (control) or 1% ethanol for 7 days. Young rice panicles were collected from five individual rice plants one day later after the treatment and were subjected to cDNA-AFLP analysis using *Eco*RI (E) and *Mse*I (M) primer combinations. The PCR products were separated in 6% denaturing polyacrylamide gel. The bands were visualized by silver staining. Solid and dashed arrows indicate the up-regulated and down-regulated TDFs, respectively. C: control and 1: 1% ethanol.

Table 3-3 Homology of differential transcript-derived fragments (TDFs) obtained from cDNA-AFLP analysis of young rice panicles treated with 1% ethanol.

TDF	Size (bp)	Expression pattern		Locus ID[a]	Description	*E*-value
		up-regulated	down-regulated			
Metabolism						
TDF 2	177	+		*Os08g0478800*	glucose-6-phosphate isomerase	3.2e-90
TDF 25	83		+	*Os03g0401300*	sucrose synthase 1	1.1e-27
TDF 31	178		+	*Os09g0509000*	DAG protein, chloroplast precursor	7.5e-88
TDF 10	103	+		*Os03g0359100*	aminotransferase domain containing protein	4.0e-43
TDF 39	242	+		*Os03g0661900*	expressed protein	3.8e-136
TDF 110	129	+		*Os01g0339900*	protein disulfide isomerase-like 2-2	7.9e-60
TDF 112	170	+		*Os02g0687900*	putative serine carboxypeptidase	5.1e-73
TDF 115	148	+		*Os06g0646400*	putative ethylene-inducible CTR1-like protein kinase	7.5e-68
Biosynthesis						
TDF 34	84		+	*Os04g0671700*	class I glutamine amidotransferase-like superfamily protein	1.0e-33
TDF 57	150	+		*Os03g0268300*	digalactosyldiacylglycerol synthase 2	1.0e-74
TDF 64	99	+		*Os02g0180500*	putative chorismatemutase	1.1e-23
TDF 66	250	+		*Os03g0245100*	probable 4-hydroxy-tetrahydrodipicolinate reductase 2	4.1e-139
TDF 100	274	+		*Os10g0518200*	cytochrome b5	7.9e-140

Table 3-3 Homology of differential transcript-derived fragments (TDFs) obtained from cDNA-AFLP analysis of young rice panicles treated with 1% ethanol. (*continued*)

TDF	Size (bp)	Expression pattern		Locus ID[a]	Description	*E*-value
		up-regulated	down-regulated			
Signal transduction						
TDF 8	130	+		*Os03g0710800*	14-3-3-like protein GF14-F	1.3e-61
TDF 37	99	+		*Os06g0478600*	leucine-rich repeat family protein-like	2.5e-41
TDF 101	179	+		*Os03g0207250*	mitochondrial ribosomal protein L51/S25/CI-B8 family protein	8.5e-60
TDF 109	163	+		*Os02g0759700*	putative F-box containing protein TIR1 1	7.9e-88
Transcriptional regulation						
TDF 44	100	+		*Os07g0497000*	expressed protein	4.4e-23
TDF 95	126	+		*Os07g0627300*	myb-related protein-like	4.8e-55
Stress response						
TDF 47	151	+		*Os07g0240300*	chloroplast lumen common protein family-like protein	1.9e-21
TDF 108	166	+		*Os07g0645100*	phosphatidylinositol 3- and 4-kinase family protein with FAT domain	2.4e-28
Membrane transport						
TDF 14	87	+		*Os02g0175700*	synaptosomal-associated protein 91-like contains ENTH (epsin N-terminal homology) domain	2.1e-31

Table 3-3 Homology of differential transcript-derived fragments (TDFs) obtained from cDNA-AFLP analysis of young rice panicles treated with 1% ethanol. (*continued*)

TDF	Size (bp)	Expression pattern		Locus ID[a]	Description	*E*-value
		up-regulated	down-regulated			
Cell cycle						
TDF 116	144	+		*Os03g0212400*	SNARE domain containing protein	2.2e-69
Unknown function						
TDF 3	136	+		*Os02g0560450*	putative uncharacterized protein	1.2e-64
TDF 5	127	+		*Os08g0179900*	putative uncharacterized protein	1.3e-61
TDF 26	91	+		*Os05g0392100*	putative uncharacterized protein	4.6e-38
TDF 27	162		+	*Os03g0594700*	expressed protein	2.0e-44
TDF 29	104	+		*Os08g0505200*	putative uncharacterized protein	5.6e-41
TDF 30	157		+	*Os01g0656600*	myosin heavy chain-like	4.0e-71
TDF 72	143	+		*Os07g0462200*	putative uncharacterized protein	1.2e-36
TDF 90	206	+		*Os02g0789100*	putative uncharacterized protein	6.4e-104
TDF 93	101	+		*Os03g0569000*	putative uncharacterized protein	1.5e-28
TDF 104	241	+		*Os04g0644800*	putative uncharacterized protein	4.4e-44
TDF 105	104	+		*Os12g0624700*	putative uncharacterized protein	1.8e-33

[a]Homology of TDF sequences to gene sequences in the Gramene rice database by BLAST search.

"+" indicates regulation pattern of TDF when sprayed 1% ethanol to the rice plants.

3.3.3 Expression analysis of ethanol-inducible genes

To confirm the expression of the ethanol-responsive TDFs obtained from cDNA-AFLP analysis, the semi-quantitative RT-PCR was performed using total RNA extracted from the control and 1% ethanol-treated young panicles as templates. Primers specific to the 34 TDFs were designed based on their homologous genes in the Gramene rice database. Furthermore, genes involved in the ethanolic fermentation pathway in rice were also examined including *alcohol dehydrogenase 1* (*adh1*), *alcohol dehydrogenase 2* (*adh2*), and *aldehyde dehydrogenase* (Lasanthi-Kudahettige et al. 2007). Three independent experiments were performed. The expression level of the rice *act1* gene was used as an internal control. After the amplification, 14 TDFs, *adh2* and *aldehyde dehydrogenase* were successfully amplified. Five of the fourteen TDFs and *adh2* showed differential expression by semi-quantitative RT-PCR. These genes showed higher expression with 1% ethanol application compared to control (**Figure 3-5**). The sequences of the 5 TDFs were homologous to a chloroplast lumen protein (*Os07g0240300*), a synaptosomal-associated protein (*Os02g0175700*), a putative uncharacterized protein (*Os05g0392100* and *Os03g0569000*), and a Myb-related protein-like *Os07g0627300* (**Table 3-3**). Information from the rice gene databases (RiceXpro database, http://ricexpro.dna.affrc.go.jp and Gramene, http://www.gramene.org) indicated that *Myb-related protein-like* (*Os07g0627300*) is expressed in several tissues of rice plants and is highly conserved across species. The promoter of this gene could be used to regulate a target gene in several tissues and in a wide range of plant species. Therefore, we focused our further study on this gene.

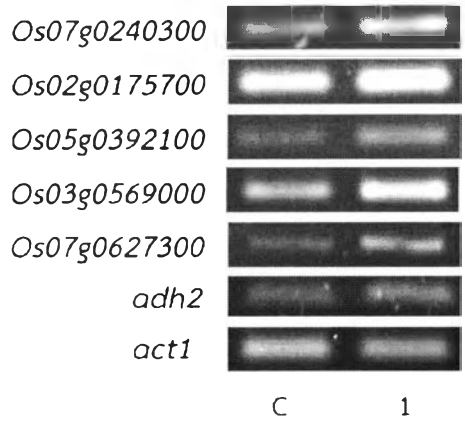

Figure 3-5 Expression analysis by semi-quantitative reverse transcription PCR of ethanol-inducible genes in young rice panicles. One-hundred and sixteen-day-old rice plants at the tillering stage prior to panicle initiation were sprayed daily with tap water (control) or 1% ethanol for 7 days. Young rice panicles were collected one day later and were subjected to expression analysis using semi-quantitative RT-PCR. The expression level of *act1* was used as the internal control. The PCR products were separated in 1% agarose gel and were visualized with ethidium bromide. All pictures are representative of three independent experiments. C: control and 1: 1% ethanol.

3.3.4 Expression analysis of *Myb-related protein-like* (*Os07g0627300*) in different tissues of rice plant

To determine expression patterns of *Myb-related protein-like* (*Os07g0627300*) in different tissues of PTT1 rice cultivars, total RNA was isolated from callus, 15-day-old seedlings, roots, stems, leaf sheaths, leaf blades, young panicles, and anthers of PTT1 rice plants grown under normal condition. The expression level of the rice *act1* gene was used as an internal control. Three independent experiments were performed. The results showed that *Myb-related protein-like* (*Os07g0627300*) was expressed in all tested rice tissues, with higher levels of expression in callus, seedlings, leaf sheaths, and leaf blades (**Figure 3-6**).

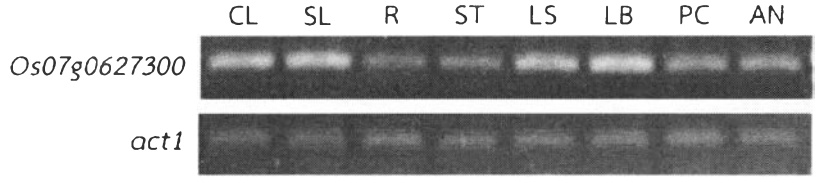

Figure 3-6 Expression analysis of *Myb-related protein-like* (*Os07g0627300*) in different tissues of rice plants. The rice plants were grown under normal condition. All samples were collected from rice plants at the reproductive stage. Semi-quantitative RT-PCR was performed with gene specific primers. Expression of *act1* was used as the internal control. The PCR products were separated in 1% agarose gel and were visualized with ethicium bromide. Similar results were obtained in three independent experiments. CL: callus, SL: 15-day-old seedling, R: root, ST: stem, LS: leaf sheath, LB: leaf blade, PC: panicle, and AN: anther.

3.3.5 Phylogenetic analysis of *Myb-related protein-like* (*Os07g0627300*)

Information from the NCBI database indicated that *Myb-related protein-like* (*Os07g0627300*) is highly conserved across species, from human to plants. To understand the evolutionary relationships of *Myb-related protein-like* (*Os07g0627300*) in rice and other species across kingdoms, phylogenetic analysis was constructed, using deduced amino acid sequences of this gene in diverse species such as *Japonica* rice, *Indica* rice, sorghum, maize, wheat, grape, Arabidopsis, green algae, frog, and human. These sequences were obtained from the NCBI Reference Sequence database (http://www.ncbi.nlm.nih.gov/refseq/). This analysis showed the phylogenetic relationship between rice *Myb-related protein-like* and the other *Myb-related protein-like* from different species. The results showed 3 main groups of

genes from human and frog, single-cell algae, and plants. In the plant group, there were two subgroups corresponding to monocot and dicot groups. As expected, *Myb-related protein-like* (*Os07g0627300*) from *Japonica* and *Indica* rice were paired together in monocot group (**Figure 3-7**).

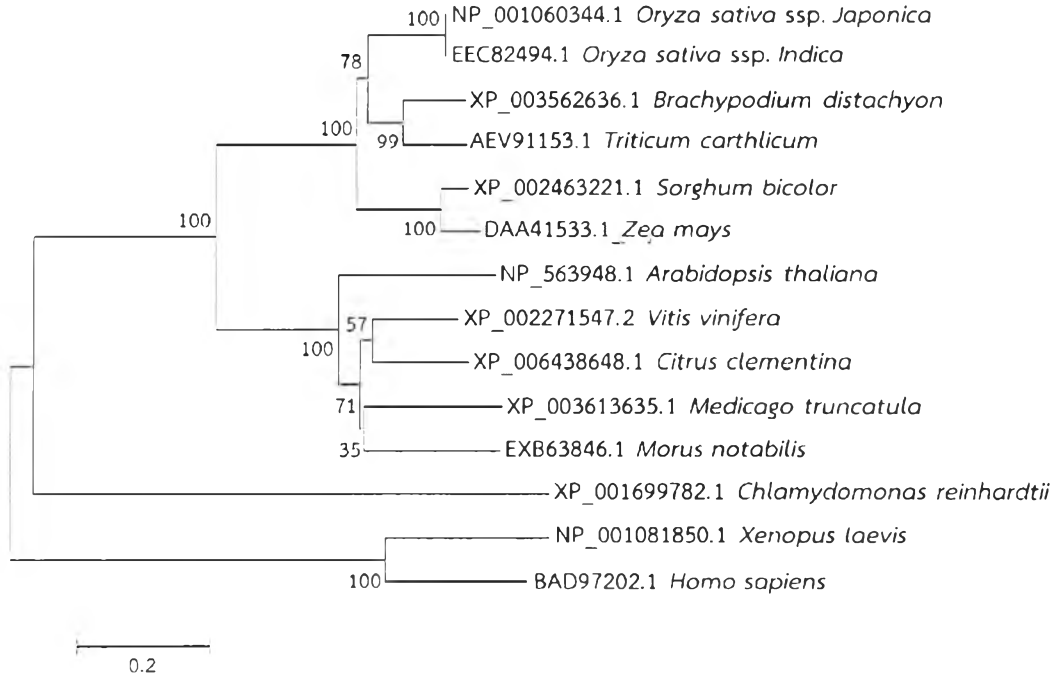

Figure 3-7 Phylogenetic relationship of rice Myb-related protein-like (Os07g0627300) and other similar sequences from different species across kingdom. A rooted neighbor-joining phylogenetic tree was constructed from the protein sequences obtained from the NCBI Reference Sequence database. Multiple sequence alignment was performed using the Clustal W in MEGA5 and the tree was generated using MEGA5. The numbers for interior branches indicate the bootstrap values (%) for 1000 replications.

3.3.6 Isolation and sequence analysis of the 5' upstream region of *Myb-related protein-like* (*Os07g0627300*)

To study sequences containing a putative promoter of *Myb-related protein-like* (*Os07g0627300*) gene , about 2,000 bp-DNA sequences upstream of the first ATG of *Myb-related protein-like* (*Os07g0627300*) from *Japonica* and *Indica* rice were obtained from the Gramene rice databases (http://www.gramene.org). and alignment of these two sequences was performed. The results showed that the 5' UTR of this gene is highly conserved. Thus, primers specific for the 5' UTR of *Myb-related protein-like* gene were designed based on the conserved sequence of *Japonica* and *Indica* rice (**Table 3-3**), and were used to amplify the 5' UTR of this gene from the PTT1 cultivar. The obtained sequence from the PTT1 cultivar was aligned and compared with the sequences of *BGIOSGA026188* from the *Indica* rice and *Os07g0627300* from *Japonica* rice reported in Gramene (http://www.gramene.org). Results showed that the 5' UTR sequence, about 1,540 bp, of PTT1 was exactly identical to the sequence of *BGIOSGA026188*. The 14-bp insertion and nucleotide substitutions were found in the sequence of PTT1 and *BGIOSGA026188* when compared with the sequence of *Joponica* rice (*Os07g0627300*) (**Figure 3-8**).

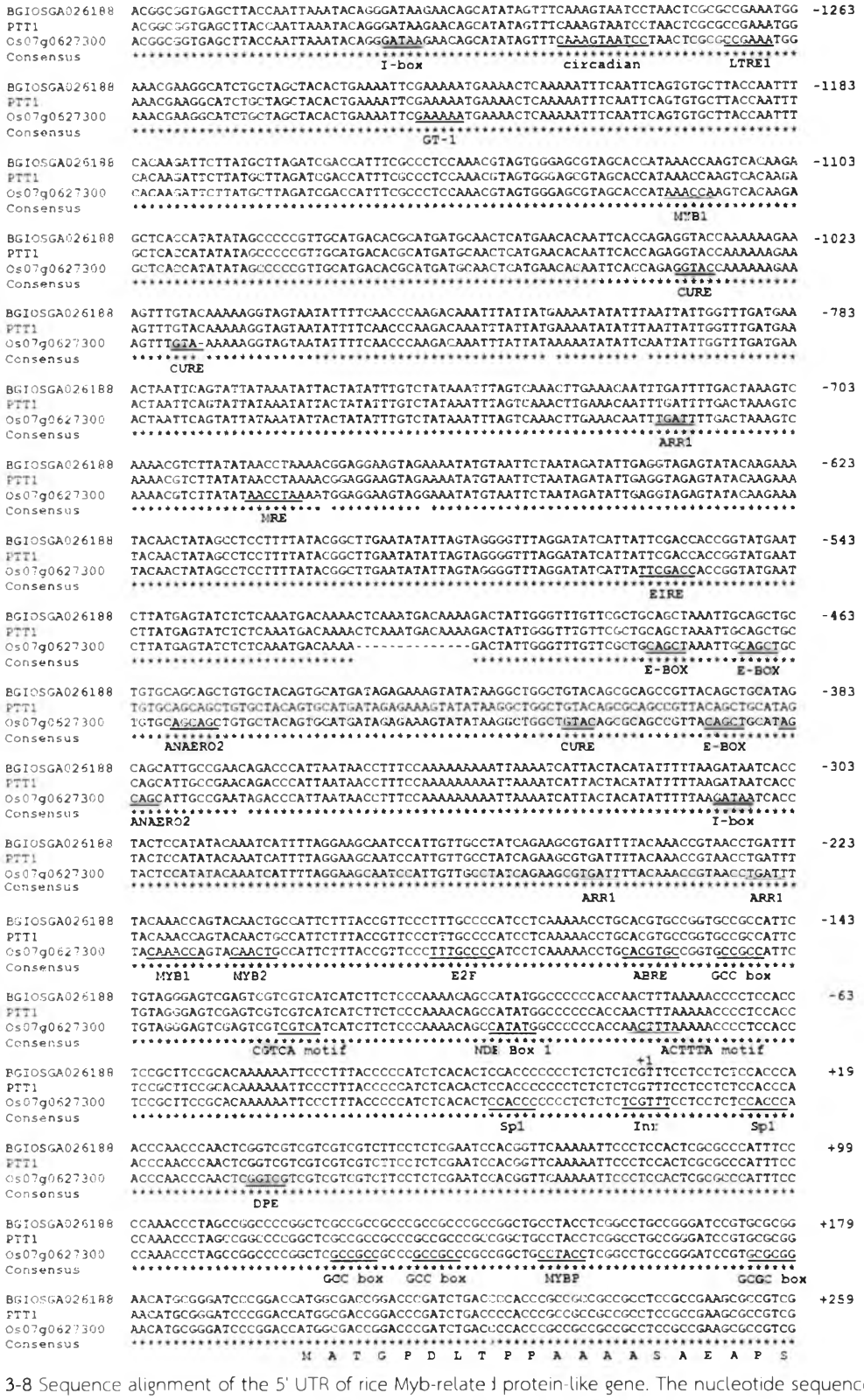

Figure 3-8 Sequence alignment of the 5' UTR of rice Myb-related protein-like gene. The nucleotide sequence of the 5' UTR of rice myb-related protein-like gene isolated from the PTT1 cultivar was compared with those of *BGIOSGA026188* from *Indica* rice and *Os07g0627300* from *Japonica* rice obtained from Gramene. Stars indicate identical nucleotides. The predicted transcription start site is indicated by +1. The ATG start site is at the position +199 relative to the TSS. The underlined sequences indicate the *cis*-acting regulatory elements.

Using the PLACE (https://sogo.dna.affrc.go.jp/cgi-bin/sogo.cgi) and PlantCARE databases (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), *cis*-acting regulatory elements in the 5' UTR sequence isolated from the PTT1 cultivar were identified. Putative *cis*-acting regulatory elements in the 1,540 bp-fragment were shown in **Table 3-4**. The obtained information indicated that the 1,540 bp-fragment of the PTT1 cultivar was a TATA-less promoter. The promoter of rice *Myb-related protein-like* gene has been annotated as a TATA-less promoter (OJ1339_F05.140) in the PlantProm DB database (http://linux1.softberry.com/data/plantprom/Links/PLPR_predicted_OS_nsite.res). The transcription start site (TSS) is located at 199 bp upstream from the translation start site (ATG). The position of TSS was designated as +1 (**Figure 3-8**). The initiator or Inr (TYRKTY) element was found at positions -2 to +4 relative to TSS. Inr functions as the TATA box in a TATA-less promoter and recognized by TFIID (Porto et al. 2014). The downstream promoter element or DPE (RGWYV) located downstream from TSS at position +34 to +38. Both Inr and DPE in TATA-less promoters are essential for transcriptional regulation (Heintzman and Ren 2007).

Four MYB recognition elements such as MYBP, MYB1, MYB2, and MRE were found. MYBP, MYB1, and MYB2 are involved in the dehydration and abscisic acid (ABA) responses (Abe et al. 2003; Grotewold et al. 1994) while MRE is involved in light response (Feldbrugge et al. 1997). The CGCG box, which is involved in multiple signaling pathways in plants (Yang and Poovaiah 2002), is located downstream from TSS. A number of the light responsive elements including Sp1 and I-box were found with differential occurrence in the obtained sequence. The ethylene-responsive element, GCC-box, has been identified in many plant species (Chakravarthy et al.

2003). This element is located in both upstream and downstream sequences from TSS of the upstream region of rice Myb-related protein-like gene. Another hormone-responsive elements including NDE box 1 in auxin response (Xu et al. 1997), CGTCA-motif in methyl jasmonate (MeJA) response (Rouster et al. 1997), and ABRE in ABA response (Kaplan et al. 2006) were also identified.

Three copies of ARR1 element are located upstream from TSS in the obtained sequence. ARR1 functions as transcriptional regulators of cytokinin-induced gene expression in plants (Ross et al. 2004). The E2F binding site was also found. The E2F transcription factors regulate transcription of genes required in DNA replication (Vandepoele et al. 2005). The E-box is another element involved in transcriptional regulation found in the 5' UTR sequence isolated from the PTT1 cultivar. The E-box is recognized by the basic helix-loop-helix (bHLH) DNA binding/dimerisation domain of transcriptional activators (Hartmann et al. 2005).

Three copper-responsive elements (CUREs) were distributed in the isolated sequence. This element is present in the 5' UTR of the copper-responsive genes that are induced under copper- and oxygen-deficient conditions (Quinn et al. 2002). The GT-1 motif was identified and located upstream from TSS. The GT-1 motif is a binding site of nuclear factors that are involved in the pathogen and salt response (Park et al. 2004). The elicitor-responsive element or EIRE was also found. The EIRE is involved in expression of genes responding to biotic and abiotic elicitors in plants (Fukuda and Shinshi 1994).

The low-temperature-responsive element or LTRE1 was found in the obtained sequence. The LTRE1 is commonly present in the promoter regions of the low-temperature-responsive genes (Dunn et al. 1998). Two ANAERO2 elements were

found and located upstream to TSS. ANAERO2 has been found in the promoter region of the anaerobic-induced genes in anaerobic tolerant plants such as rice, maize, and Arabidopsis (Mohanty et al. 2005). The circadian-responsive element was also identified.

Table 3-4 Predicted *cis*-acting regulatory elements found in the Os07g0627300 promoter by PLACE[a] and PlantCARE[b] databases

Element	Sequence[c]	Function	Position relative to TSS[d]	Reference
CGCG box	VCGCGB	AtSR binding site, AtSRs involved in multiple signaling pathways in plants	+174	Yang and Poovaiah (2002)
MYBPZM	CCWACC	maize P (myb homolog) binding site	+149	Grotewold *et al.* (1994)
GCC-box	GCCGCC	ethylene-responsive element, JA-responsive gene expression, defence-related gene expression	+134, +124, -151	Chakravarthy *et al.* (2003)
DPE	RGWYV	a posterior TFIID binding site	+34	Kutach and Kadonaga (2000)
Sp1	CCRCCC	involved in the light responsiveness	+13, -18	Thanh *et al.* (2012)
Inr	TYRKTY	TFIID binding site	-2	Javahery *et al.* (1994)
TCCACCT-motif	TCCACCT	-	-67	Nalbandi *et al.* (2013)
ACTTTA-motif	ACTTTA	BtBBF1 binding site, required for tissue-specific and auxin-regulated expression of the *rolB* oncogene	-81	Baumann *et al.* (1999)
NDE Box 1	CATATG	NDE binding site, involved in auxin responsiveness	-98	Xu *et al.* (1997)
CGTCA-motif	CGTCA	involved in the MeJA-responsiveness	-123	Rouster *et al.* (1997)
ABRERATCAL	MACGYGB	involved in the ABA responsiveness	-162	Kaplan *et al.* (2006)
E2FCONSENSUS	WTTSSCSS	E2F-binding site	-185	Vandepoele *et al.* (2005)
MYB2CONSENSUSAT	YAACKG	MYB recognition site, involved in the dehydration and abscisic acid responsiveness	-209	Abe *et al.* (2003)
MYB1AT	WAACCA	MYB recognition site, involved in the dehydration and abscisic acid responsiveness	-218, -1117	Abe *et al.* (2003)
ARR1AT	NGATT	cytokinin-regulated ARR1 binding site, involved in transcription	-227, -247, -718	Ross *et al.* (2004)
IBOXCORE	GATAA	involved in the light responsiveness	-311, -1311	Terzaghi and Cashmore (1995)
ANAERO2CONSENSUS	AGCAGC	involved in the fermentative pathway	-383, -456	Mohanty *et al.* (2005)
EBOXBNNAPA	CANNTG	regulated the transcription of genes	-392, -468, -479	Hartmann *et al.* (2005)
CURECORECR	GTAC	copper-response element, involved in the oxygen responsiveness	-409, -856, -1035	Quinn *et al.* (2002)
EIRE	TTCGACC	required for the elicitor responsiveness	-560	Fukuda and Shinshi (1994)
MRE	AACCTAA	MYB recognition site, involved in the light responsiveness	-209	Feldbrügge *et al.* (1997)
GT1GMSCAM4	GAAAAA	involved in the pathogen and salt responsiveness	-1227	Park *et al.* (2004)
LTRE1HVBLT49	CCGAAA	low-temperature-responsive element	-1270	Dunn *et al.* (1998)
CIACADIANLELHC	CAANNNNATC	involved in the circadian expression	-1290	Piechulla *et al.* (1998)

[a]PLACE website http://www.dna.affrc.go.jp/cgi-bin/sogo.cgi; [b]PlantCARE website http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

[c]symbols used in the sequences are listed as follow. R: G or A, Y: T or C, M: A or C, K: G or T, S: G or C, W: A or T, B: G or T or C, V: G or C or A, and N: G or A or T or C.

3.4 Discussion

An ethanol-inducible system or '*alc* switch' is a useful tool for gene regulation and expression studies (Tomsett et al. 2004). In the presence or absence of ethanol, the *alc* switch regulates expression of target genes in spatial and temporal patterns. This advantage could diminish undesirable effects from overexpression of interested genes in the tested plants (Li et al. 2005). However this ethanol-inducible system is derived from the ascomycete fungus, *Aspergillus nidulans* (Felenbok et al. 1988). *A. nidulans* has a substantial genetic distance from plants, thus the use of the ethanol-inducible system and genes derived from plant itself would be of interest. The responses of non-transformed plants including potato (Vreugdenhil *et al.* 2006) and sugarcane (Camargo et al. 2007) to ethanol application indicated that higher plants including rice may have both putative ethanol-inducible promoters and ethanol-regulated transcription factors. Thus, it is possible to isolate only the ethanol-inducible promoter from rice plants and use it to control target genes for functional analysis.

In this study, ethanol-inducible genes in young panicles of ethanol-treated rice plants were identified using cDNA-AFLP. The expression of ethanol-inducible genes was confirmed by semi-quantitative RT-PCR. Moreover, the 5' UTR of a highly conserved gene was isolated and analyzed for *cis*-acting regulatory elements. Our aims were to use a minimal concentration of ethanol that could induce gene expression but is non-toxic to the treated plants. To attain this goal, ethanol at various concentrations were applied to rice plants at the seedling and reproductive stages, and its effects on growth of rice plants were determined at the seedling, reproductive, and ripening stages. The results showed that ethanol at high

concentration affected growth of rice plants. Shoot dry weight and shoot length were significantly decreased with the application of ethanol higher than 1% (**Figure 3-1**). Similarly, application of ethanol higher than 1% affected the elongation and dry weight of coleoptile and root in four-day-old rice seedlings (Kato-Noguchi and Kugimiya 2001). The application of 2% ethanol to rice plants at the reproductive stage affected growth and several agronomic traits including grain yield (**Figures 3-2 and 3-3**).

It has been reported that ethanol at the concentration of 2% or lower did not affect growth of plants (Ait-ali et al. 2003); however, ethanol at high concentration could cause dehydration (Camargo et al. 2007). The effects of ethanol on rice growth probably depend on the amount of ethanol used, duration of time of the treatment, and developmental stages of the plant. In addition, different plant species may respond differently to ethanol application. In this study, the amount of ethanol and time duration used for rice plants were greater than that used for 3-week-old Arabidopsis seedlings (Ait-ali et al. 2003) and 40-day-old sugarcane plantlet (Camargo et al. 2007). However, in general our results were in accordance with their results (Ait-ali et al. 2003; Camargo et al. 2007). Our results showed that 1% ethanol did not affect rice growth. Therefore, 1% ethanol was used to induce the expression of ethanol-inducible genes in young rice panicles.

cDNA-AFLP analysis was used to identify ethanol-inducible genes in young panicles from ethanol-treated rice plants. cDNA-AFLP is a useful tool for genome-wide expression analysis (Breyne et al. 2003). This method has been used to isolate and identify genes that respond to stress in many plants species such as rice (Ventelon-Debout et al. 2008), sugarcane (Camargo et al. 2007), and tomato (Nazeem

et al. 2011). In this study, immobilized cDNA was used in double digestion. The use of immobilized cDNA can minimize the redundancy of cDNA fragments when using more than one restriction enzyme in digestion (Weiberg et al. 2008). Thirty-four ethanol-responsive genes were obtained from our cDNA-AFLP analysis (**Figure 3-4 and Table 3-3**), and expressions of some of these ethanol-responsive genes were confirmed by semi-quantitative RT-PCR. In addition, genes involved in the ethanolic fermentation pathway in rice (Lasanthi-Kudahettige et al. 2007) including *adh1* (*Os11g0210300*), *adh2* (*Os11g0210500*), and *acetaldehyde dehydrogenase* (*Os05g0536400*) were also examined by semi-quantitative RT-PCR.

Our results showed that *chloroplast lumen protein* (*Os07g0240300*), *synaptosomal-associated protein* (*Os02g0175700*), *putative uncharacterized protein* (*Os05g0392100* and *Os03g0569000*), and *Myb-related protein-like* (*Os07g0627300*), and *adh2* in young rice panicles responded to exogenously added ethanol (**Figure 3-5**). The gene encoding a chloroplast lumen protein (*Os07g0240300*) and *adh2* (*Os11g0210500*) were involved in stress responses. The chloroplast lumen protein (*Os07g0240300*) was reported to be involved in the oxidative stress response. Chloroplast lumen proteins protect lipids in the thylakoid membrane from oxidative stresses (Levesque-Tremblay et al. 2009). An ethanol stress may cause the production of reactive oxygen species (ROS) within plant cells, resulting in up-regulation of this gene (Hirayama et al. 2004). *adh2* (*Os11g0210500*) encoding alcohol dehydrogenase 2 (ADH2) is involved in the ethanolic fermentation pathway. ADH2 converts ethanol to acetaldehyde which is then converted to acetate by acetaldehyde dehydrogenase (ALDH). This results in anaerobic tolerance in rice (Hirayama et al. 2004; Lasanthi-Kudahettige et al. 2007). The excessive application of

ethanol may induce the *adh2* expression by altering an endogenous ethanol level in rice plants. The gene encoding synaptosomal-associated protein (*Os02g0175700*) was up-regulated under ethanol induction. Synaptosomal-associated protein is involved in the clathrin coat assembly which plays an important part in signal transduction and in plants. The clathrin-mediated endocytosis was stimulated by ROS produced in plant cells (Leborgne-Castel et al. 2008). Therefore, it is possible that ROS produced from an ethanol stress may activate this gene function. Two genes of unknown functions, *Os05g0392100* and *Os03g0569000*, also responded to ethanol. Ethanol may up-regulate these genes via one or more of pathways in rice plants. *Os07g0627300* encoded Myb-related protein-like involved in transcriptional regulation. Under stress conditions including dehydration, this protein regulates the expression of genes involving different pathways in plants (Xie et al. 2010b). Excessively applied ethanol may cause a dehydration stress (Camargo et al. 2007) that induces the transcription of this gene in rice plants.

Among these 6 ethanol-inducible genes, *Os07g0627300* encoding for Myb-related protein-like has a low basal expression (**Figure 3-5**). Moreover, the information from the NCBI database indicated that the amino acid sequence of this gene is highly conserved during evolution. Our phylogenetic analysis reveals the relationship of conserved Myb-related proteins among plant, algae, animals, and humans. Myb-related proteins from *Indica* (*BGIOSGA026188*) and *Japonica* (*Os07g0627300*) were closed together in a monocot subgroup of the plant group (**Figure 3-7**). *Os07g0627300* is an orthologue of *MYB88* in Arabidopsis, whose paralogue is *FOUR LIPS* (*FLP*). FLP/MYB88 is a transcription factor that regulates the expression of many genes involved in different pathways (Xie et al. 2010a).

FLP/MYB88 has been found to play an important role in abiotic stress response in plants (Xie et al. 2010b). Under abiotic stress, expression of genes involved in cell division of stomata was down-regulated by FLP/MYB88 (Xie et al. 2010a) while stress-responsive genes were up-regulated (Xie et al. 2010b). This information suggested that *FLP/MYB88* could regulate genes across different pathways in plants including cell development, signal transduction, and stress response. Furthermore *Myb-related protein-like* (*Os07g0627300*) has been reported to be expressed in several tissues of rice plants (RiceXpro database). Consistent with the expression pattern of *Os07g0627300* reported in the RiceXpro database, our results showed that *Os07g0627300* was expressed in callus, seedlings, and several tissues including roots, stem, leaf sheaths, leaf blades, young panicles, and anther of Pathumthani 1 (PTT1) rice cultivars (**Figure 3-6**). Based on the conservation, functions, and expression pattern of *Myb-related protein-like* (*Os07g0627300*), the promoter of this gene could be used to regulate expression of genes in a wide range of species and tissues.

The 5' UTR of *Myb-related protein-like* (*Os07g0627300*) was isolated from genomic DNA of *Indica* rice cultivar, PTT1. The 1,540 bp DNA fragment was obtained. Results from sequence alignment indicated that the 5' UTR of this gene is highly conserved among *Indica* and *Japonica* rice, and this fragment from the PTT1 cultivar is identical to that in *BGIOSGA026188* from *Indico* rice (**Figure 3-8**). The *cis*-acting regulatory elements in a 1,540 bp-fragment were identified using the PLACE and PlantCARE databases (**Figure 3-8 and Table 3-4**). In accordance with the information from the PlantProm DB database (http://linux1.softberry.com/data/plantprom/Links/PLPR_predicted_OS_nsite.res), the 5' UTR of *Myb-related protein-like* (*Os07g0627300*) isolated from the PTT1 cultivar was a TATA-less promoter. This type of promoters

has been found to regulate expression of some large groups of genes such as photosynthesis genes, genes encoding growth factors, and housekeeping genes (Smale 1997; Solovyev et al. 2010). The putative *cis*-acting elements required to initiate transcription by a TATA-less promoter are the initiator (Inr) and the downstream promoter element (DPE) (Heintzman and Ren 2007; Smale 1997). In this study, both Inr and DPE were found in the obtained 5' UTR, and the transcription start site (TSS) is located in Inr. The consensus sequence of Inr was $TCG_{+1}TTT$ which is consistent with the YR rule in rice Inr sequence, where $G_{+1}$is TSS. A dimer motif of Y (C/T) and R_{+1}(A/G) or the YR rule has been identified at the TSS (R_{+1}) in rice and Arabidopsis (Yamamoto et al., 2007). The DPE is generally located at positions +28 to +32 to the TSS (Kutach and Kadonaga 2000; Porto et al. 2014). In this study, we found that DPE was located at position +34 to +38 to the TSS. This Myb-related gene is orthologous to *FLP/MYB88* which senses stress conditions (Xie et al. 2010b).

Concordantly, a number of motifs involved in biotic and abiotic stress responses were identified in the 5' UTR of *Myb-related protein-like* (*Os07g0627300*). These elements differentially occurred in the obtained sequence. Several elements responding to dehydration and ABA including MYB1, MYB2, and ABRE were found. These elements may contribute to up-regulation of *Myb-related protein-like* (*Os07g0627300*) when ethanol was applied and induced dehydration in rice plants. The presence of CUREs and ANAERO2 in the obtained sequence suggested that the anaerobic condition could induce expression of this gene. Moreover, elements responding to light, hormone, and elicitors were also identified in the 5' UTR of *Myb-related protein-like* (*Os07g0627300*).

3.5 Conclusion

This study reported the response of rice plants to exogenously added ethanol at phenotypic and molecular levels. Six of ethanol-inducible genes in young rice panicles were identified. The 5' UTR of highly conserved *Os07g0627300* encoding Myb-related protein-like was isolated and analyzed for *cis*-acting regulatory elements. Our results suggested that *Myb-related protein-like* (*Os07g0627300*) may be involved in multiple response systems in rice plants by controlling expression of genes involved in several pathways. The promoter of this gene could be used to regulate expression of numerous target genes in various plant tissues. The obtained results could provide a useful tool to regulate agronomic desirable genes in cereal crop plants. Furthermore, the ethanol-inducible promoter of this gene may be used for functional genomic analysis of specialty rice cultivars including antioxidant-rich and stress-tolerant rice. The obtained knowledge would be useful for the development and production of new hybrid rice cultivars.