

Transcript abundance of *rml1*, encoding a putative GT1-like factor in rice, is up-regulated by *Magnaporthe grisea* and down-regulated by light

Rong Wang^{a,b}, Guofan Hong^{a,b}, Bin Han^{a,*}

^aNational Center for Gene Research, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences,

500 Caobao Road, Shanghai 200233, China

^bShanghai Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences,

320 Yueyang Road, Shanghai 200031, China

Received 28 March 2003; received in revised form 8 August 2003; accepted 5 September 2003

Received by G. Theissen

Abstract

We isolated and sequenced both genomic DNA and cDNA clones, which encoded a putative GT1-like protein with 385 amino acids, from cultivated rice (*Oryza sativa* ssp. *indica*). This protein shows significant amino acid sequence similarities with trihelix DNA-binding GT-1a/B2F and GT-1 factors that were identified in dicot plants. Northern blotting analysis indicated that the transcript of the rice GT-1 factor in seedling was up-regulated by the rice blast fungus *Magnaporthe grisea*, down-regulated by various continuous light conditions and expressed rhythmically in light/dark cycles. This GT1-like factor gene was therefore designated as *rml1* (rice gene regulated by *M. grisea* and light). The putative RML1 protein, encoded by this single copy gene, is thus identified as a new member of the plant-specific GT family of transcription factors in rice.

© 2003 Elsevier B.V. All rights reserved.

Keywords: *Oryza sativa indica*; Phytochrome; *Magnaporthe grisea*; GT1-like DNA binding protein

1. Introduction

To ensure optimal and healthy growth, plants adopt a variety of means to respond to environment cues such as light, abiotic agents, microbes, etc. Such responses are involved in regulating transcriptional processes. Several classes of *cis*-acting sequences and their cognate proteins have been identified in higher plants under various conditions. One class of *cis*-elements is the highly degenerate GT-element (Zhou, 1999). GT-elements have both a positive and a negative function in modulating cell-type-specific transcription (Villain et al., 1996). Synthetic promoters containing GT-elements were also found to be regulated by light (Lam and Chua, 1990), although these elements have been

found in the promoters of some light-independent genes. The pathogen defense-related tobacco *PR-1a* gene contains GT-elements in the promoter region (Buchel et al., 1999).

Nuclear proteins binding to GT-elements were initially cloned from rice and tobacco (GT-2 and GT1a/B2F, respectively) (Dehesh et al., 1990; Gilmartin et al., 1992). Rice GT-2 contains two separate trihelix (helix–loop–helix–loop–helix) structures, which can bind to the GT-2 and GT-3 boxes in the rice *phyA* promoter. Tobacco GT1a/B2F has only a single trihelix domain that is responsible for specific binding to the BoxII core sequence in the ribulose 1,5-bisphosphate carboxylase/oxygenase small subunit (*rbcS*) gene promoter. GT-factors corresponding to rice GT-2 or tobacco GT-1a have been characterized from *Arabidopsis* and seem to have the same DNA-binding preferences (Kuhn et al., 1993; Hiratsuka et al., 1994). Rice GT-2 and *Arabidopsis* GT-1 have been shown to activate transcription in vivo (Ni et al., 1996; Le Gourrierec et al., 1999).

Though GT-elements can modulate both light activation and light repression, depending on the promoter context,

Abbreviations: cM, centimorgan; h, hour; *M. grisea*, *Magnaporthe grisea*; min, minute; ORF, open reading frame; *rbcS*, ribulose 1,5-bisphosphate carboxylase/oxygenase small subunit gene.

* Corresponding author. Tel.: +86-21-64845260; fax: +86-21-64825775.

E-mail address: bhan@ncgr.ac.cn (B. Han).

transcriptional regulation of the cloned trihelix factors appeared to be independent of light, with the exception of soybean *GmGT-2*, which was down-regulated by light (O'Grady et al., 2001).

Here, we cloned a new gene (*rml1*) encoding a putative GT1-like protein from rice through the genomic DNA sequencing analysis and cDNA cloning. Putative *cis*-elements were predicted in the proximal promoter sequence. Particular attention is given to the expression pattern of the *rml1* gene with respect to tissue specificity, to rice blast *Magnaporthe grisea*, to light and to rhythmical expression, which is poorly investigated until now.

2. Material and methods

2.1. BAC DNA sequencing and analyzing

A rice bacterial artificial clone (BAC) H0302E05 from the rice BAC library of *Oryza sativa indica* Guangluai 4 was anchored on the long-arm of chromosome 4 as determined using probe R896 provided by the rice genome program (RGP) in Japan. Shotgun and pairwise-end sequencing were used for BAC sequencing. The sequences were assembled with PHRED/PHRAP software and edited by GAP4 software. Sequence annotation was carried out for the BAC DNA sequence using a combination of software prediction and database searches. The sequence of BAC H0302E05 was submitted to the EMBL database, with accession number AL627350.

2.2. Rice material and treatment

Rice seeds (*O. sativa indica* Guangluai 4) were sterilized and imbibed then germinated at 37 °C. Green plants grew at 30 °C under white light with a day length of 12 h. For etiolated plants, germinated seeds grew in black containers placed in a dark environmental chamber at 30 °C. Calli induced from the scutellia were kept in the dark and subcultured every 2 weeks on solid N6 media supplemented with 3% sucrose, and 2 mg/l 2,4-dichlorophenoxyacetic acid.

Conidia of *M. grisea* were washed from filter paper with Tween-20 (0.02%) and adjusted to a concentration of 105 conidia/ml; 4-week-old rice plants were inoculated by

spraying the conidial suspension or 0.02% Tween-20 onto the leaves. Treated rice plants remained in the light.

For light regulation experiments, calli and 7-day-old etiolated seedlings were exposed to different irradiation levels for various lengths of time before harvesting. The light sources were white fluorescent tubes with filters for red, far-red or blue light. The fluence rates used were 6.36 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for red light, 10 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for far-red light and 4.24 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for blue light.

2.3. Molecular cloning of cDNA fragments

Bioinformatics approaches were used to predict the presence of new gene and the overlapping PCR strategy was used for amplification of cDNA instead of construction and screening a cDNA library. Based on the exon prediction results, we designed sets of primers for RACE and RT-PCR. The design primers allowed for the generation of overlapping cDNA fragments, thus enabling determination of the entire mRNA sequences. cDNA amplification was carried out using a series of Clontech system. Briefly, the first strand extension was preceded with PowerScript™ amplification system. Then, the 3'- and the 5' GC-rich cDNA were amplified in PCR with 'Advantage cDNA polymerase Mix' and 'Advantage-GC cDNA polymerase Mix', respectively. CDSIII/3' PCR primer (modified Oligo dT), Smart IV and 5' PCR primer come from the SMART™ cDNA construct Kit. Primers complimentary to positions of the BAC sequence and their annealing temperature are indicated before each primer sequence. The PCR conditions were selected according to the manufacturer's instruction and the primers annealing temperature. Some PCR products were sequenced directly with gene specific primer and some were cloned in the pGEM-T Easy vector (Promega) then sequenced.

We performed nested 5'-RACE to verify the 5'-ends of the *rml1* mRNA. The 5' end anchor primer Smart IV and a reverse primer R1 (3118–3152, 66.3 °C, 5'-ACCGGCA-TATCCCTGTCTAAGGAACGGACAACCTC-3') were used in the original cDNA extension. Two rounds of RACE reactions were performed with the original cDNA. The initial round of PCR was conducted with the reverse primer R1 and the 5' PCR primer (The sequence of 5' PCR primer is same as the 5'-end sequence of SmartIV). With two pairs of primers, two different fragments were synthesized in the second round of amplification. The PCR products named

Fig. 1. Analysis of genomic sequence and identification of the two transcripts of *rml1* gene. (A) Genomic organization and partial sequence of the *rml1* gene. Line numbers represent the positions of nucleotide acids sequences on the BAC clone. The protein sequences and the coding nucleotides are presented in single letter code and in triplets, respectively. The amino acids sequences of four closely spaced helices in the DNA-binding domain are underlined. Intronic sequences of *rml1-a* transcript are not shown except for short areas around the splice junctions, which are shown with italic letters. The start and stop codons of RML1 protein, the exon–intron junctions of two transcripts are also underlined. The polyadenylation signal element (AATAAA) and site (CA) are in boldface. The 5' exon sequences of *rml1-b* transcript are shown in red letter. The start and stop sites of two probe (Prml1 and Prml2), which are used in Northern blot analysis, are indicated with two arrows separately. For full genomic sequence, see our submission #AL627350. (B) RNA gel analysis of the two transcripts. RNA from calli (Lane C) or etiolated seedling leaves (Lane EL) was hybridized with both Prml1 and Prml2 probes mixtures. (C) Alignment of the amino acid sequences of the DNA binding domains of rice RML1, *Arabidopsis* GT1 (GT1-A, accession number AAA66473) and tobacco GT1a/B2F (GT1-T, accession number AAA34085) by the GCG Pileup program. RML1 displays identity of 65% and 62% to GT1 and GT1a/B2F over the triplex DNA binding domain. The predicted four-helix region is indicated in red.

RML1-5' was amplified with a reverse primer R2 (2409–2425, 52.2 °C, 5'-AGTGGATGACTATCAGG-3') and a forward primer F1 (23–40, 51.1 °C, 5'-CTGACCCAC-TAGTCCATG-3'). The products called RML2-5' were with a reverse primer R3 (3082–3113, 65 °C, 5'-TCCTCATCTTCAAGCCAAAAGGCCCGTCTTTGTTC-3') and the forward primer 5' PCR primer. For 3' RACE, the first-strand cDNA was synthesized using the CDSIII primer and separate RNA pools prepared from root, leaf and callus at different times. To obtain specific 3' RACE product, named as RML3', two gene-specific forward primers (F2: 640–656, 54.6 °C, 5'-AGATGGCGTGCTA-CAAG-3'; F3: 2406–2422, 52.2 °C, 5'-GAACCTGA-TAGTCATCC-3') were used with the CDSIII primer in two rounds of reaction.

A fragment RML1m, which overlaps both RML1-5' and RML3' fragments, was synthesized with a reverse primer R4 (4129–4145, 54.6 °C, 5'-GTAGCCTGAATACTCCC-3') and a forward primer F2. We constructed *rm1_a* and *rm1_b* cDNA as shown in Fig. 2A. These sequences were submitted to EMBL database, with accession numbers AJ535489 and AJ535490, respectively.

2.4. Cloning of promoter region by genomic PCR

BAC (AL627350) does not contain the entire promoter region of the *rm1* gene. Based on the colinearity between the cultivated rice subspecies *japonica* and *indica* genomes (Zhao et al., 2002), we designed primers ProPst, with an incorporated *Pst*I site (3–28, 70 °C, 5'-ctcctgcag GGTCAGGCCCATCAAAATGTACAAGC-3') and ProEcoR, with an incorporated *Eco*RI site (65.5 °C, 5'-ggaattc TCTTACTGCTTTATTTCTCCCTACTGTCG-3') according to the *japonica* BAC clone (AL606998) to amplify the promoter sequence. A 1 µg *Pst*I-digested genomic DNA was used as template. After verifying that the promoter sequence generated by PCR was homologous with the corresponding *Japonica* sequence, we inserted this *Eco*RI/*Pst*I fragment into pGEM-T Easy vector.

2.5. Southern and Northern blot hybridization

Total genomic DNA was isolated in 20 ml buffer containing 7 M urea, 0.35 M NaCl, 50 mM Tris-HCl

(pH 8.0), 20 mM EDTA (pH 8.0) and 6.25% redistilled phenol. 10 µg genomic DNA digested with *Pst*I, *Xba*I or *Hind*III was electrophoresed on a 0.8% agarose gel, transferred to nylon-N⁺ membrane and probed with 1.18 kb RML3' fragment. Hybridization conditions were according to the protocol of the ECL kit (Amersham Pharmacia).

Total RNA was isolated from seedlings or calli with the RNeasy Plant Mini Kit (Qiagen) and treated with RNase-free DNase (Qiagen) to avoid genomic DNA contamination. For RNA gel blot analysis, 15 µg total RNA was separated on 0.6 M formaldehyde agarose gels, and cross-linked with a UV cross-linker (Bio-Rad). To determine the level of the two-*rm1* gene transcripts, we used two probes in Northern blot analyses. The positions of the two probes are shown in Figs. 1A and 2A. Probe Prm1, used for detecting *rm1_a* transcripts, is the PCR product obtained using a reverse primer (622–654, 68.4 °C, 5'-TCCTTGTCACGC-CATCTTGGCGGGCAATTC-3') and a forward primer (223–237, 62.3 °C, 5'-CGGCCATGCTCCTCTCC-3'). Probe Prm2, used to detect *rm1_b* transcripts, is the PCR fragment generated using a reverse primer R2 and a forward primer (1822–1841, 51.7 °C, 5'-GCAATAATTCAGATATGCC-3'). Gene-specific probes were labelled with α-³²P according to the manufacturer's instructions (Amersham rediprimer™ II labeling system). Membranes were hybridized according to the protocol of expressHyb hybridization solution (Clontech). Loading of the samples was monitored by rehybridization of the nylon membrane with a rice 18S ribosomal DNA probe.

3. Results

3.1. Molecular cloning and identification of the cDNA clones of *rm1*

Through the DNA sequencing analysis of the BAC clone H0302E05, we identified a putative new gene encoding a GT1-like protein with a single trihelix DNA binding domain (Fig. 1A). This gene was designated as *rm1* in this study. Genomic DNA sequence analysis revealed that the *rm1* gene had five exons and encoded a putative GT1-like protein with 385 amino acids (Fig. 1A). To clone the

Fig. 2. Characterization of the *rm1* gene. (A) Schematic presentation of the *rm1* gene structure. Upper and lower panels are two transcripts of the *rm1* gene, indicated by *rm1_a* and *rm1_b*. Middle panel is the genomic structure of *rm1* gene; in which, solid bars, connecting lines and dot lines respectively represent exons, introns and the elongated promoter region. The exon–intron structure was established by comparison of the cDNA with genomic sequence. The sequence surrounding the splice junction between exons 1 and 2 of *rm1_b* is indicated under the arrow and the non-canonical intron designation is indicated in bold lower case. The positions of two probes (Prm1 and Prm2) used in Northern blot analyses are indicated by gray box. The fragments used to construct the two-*rm1* cDNAs and restriction sites *Sac*II, or *Nco*I used in overlapping are indicated. Numbers after each restriction enzyme show the restriction sites complimentary to positions of the BAC sequence. (B) Promoter region and potential regulatory *cis*-elements. Numbering is from the putative start site of translation, shown with an angled arrow. Bold type letters showed the putative TATA-like and CAAT-like boxes. The 8 bp repeat sequences close to REalpha, REbeta are outlined in italic letter. Putative binding sites for transcription factors are boxed according to their position and consensus sequences found in Table 1. Arrows near the *cis*-element names indicate the orientation of the binding site. (C) Southern blot analysis of the *rm1* gene. Rice genomic DNA digested with *Hind*III (H), *Xba*I (X) or *Pst*I (P) was electrophoresed in lanes H, X, and P, respectively. The positions of λ *Hind*III size markers are indicated on the right-hand site.

Table 1
The potential *cis*-elements of the *rml1* gene

DNA-binding factor	Matrix (#1000 bp)	Consensus	<i>rml</i> sequence	Relative fit	Position
HomeoBox protein, DOF2	Dof2(2.78)	NNW <u>AAAG</u> CNNN	AATAAAGCAGT	0.996	– 1022(–)
HomeoBox protein, DOF3	Dof3(3.32)	NNN <u>AAAG</u> CNNN	CAGAAAGGATG	0.967	– 939(+)
			AAA <u>AAAG</u> ATTA	0.970	– 807(–)
			TGTAAAGTATC	0.983	– 780(–)
			GGAAAAGGGAA	0.982	– 686(+)
			GAAA <u>AAAG</u> AAGA	0.978	– 548(–)
			ATAAAAGITTT	0.973	– 427(–)
			CTAA <u>AAAG</u> CAAA	0.994	– 409(–)
			AGCAAAGGCTG	0.969	– 404(–)
			CGAAAAGATGA	0.972	– 306(+)
			TCAA <u>AAAG</u> ACAA	0.980	– 274(+)
Dof 1/MNB1a	Dof1(1.96)	NNW <u>AAAG</u> NNNN	ATTAAAGTTTA	0.987	– 378(–)
			GTTAAAGCTTG	0.984	– 230(+)
bZIP, G-/C-box protein	TGA1A(2.21)	NNSACGTSNNN	AATACGTGATG	0.944	– 857(+)
			CGAACGTCCAC	0.920	– 152(+)
	CPRF(2.12)	NNC <u>ACGT</u> GNNN	TATACGTGTTT	0.935	– 428(+)
Opaque-2	Opaque-2(1.90)	NNNNNCC <u>ACGT</u> NNNNN	GTGCATC <u>ACGT</u> ATTGGT	0.970	– 843(–)
			TAAGAAC <u>ACGT</u> ATATAA	0.950	– 415(–)
	Opaque(0.03)	NNNNNCCAT <u>TCATC</u> NN	GAGGTACACT <u>TCATC</u> CAA	0.978	– 526(–)
WRKY	W-box(0.25)	NNNST <u>TTGAC</u> CYNNNN	AATGAT <u>TTGAC</u> AAGAT	0.972	– 717(+)
RAV1	RAV3(5.36)	NGCA <u>ACAK</u> AWN	TCCA <u>ACAG</u> CGT	0.910	– 684(–)
AHBP	ATHB(0.20)	NCA <u>ATTAT</u> TNNNN	CTC <u>ATTAT</u> TTGAAC	0.894	– 887(–)
			TAA <u>ATTAT</u> TTTTT	0.949	– 445(–)
SBF-1	SBF1(0.43)	KWRTNG <u>TAA</u> WWWN	TTGGAG <u>TAA</u> TTTTTA	0.901	– 354(–)
			TTAAGG <u>TAA</u> AAATT	0.950	– 349(+)
			TATCAG <u>TAA</u> AAATT	0.916	– 321(–)
Myb-like protein	MYBPH(0.15)	NNAAA <u>CSGTT</u> AN	ATCTTGCG <u>GTTAT</u>	0.805	– 804(+)
			ACTTATCAG <u>TAA</u>	0.810	– 318(–)
Unknown	I-box(0.03)	NWWRMGATAAGRTTATN	CCTCCGATAAGCCAAT	0.843	– 485(–)
			TAACTGATAAGTATAAG	0.890	– 329(+)
Unknown	REbeta(0.24)	CGGATA	<u>CGGATA</u>	–	– 620(+)
			<u>CGGATA</u>	–	– 519(–)
Unknown	REalpha	AACCAA	<u>AACCAA</u>	–	– 598(+)

The matrixes and consensus motifs are from the Transfac database except that REbeta and REalpha elements are from PLACE database. The cores of the consensus matrixes are underlined. The number in brackets after each matrix (except REalpha) is the frequency that is possible to be detected in a random sequence of 1000 nucleotides. The plus or minus in brackets after each position number indicated the same or reversal orientation of the binding site with that of the translation of *rml1*, respectively.

expressed *rml1* gene, RT-PCR and RACE reaction were performed using primers located in different predicted exons. We initially cloned the cDNA corresponding to this protein by RACE. DNA sequence analysis revealed that the transcript of RML1-5' with 718 bp in length ended before a putative ATG initiation codon and was found to be identical to the 5' end of the predicted coding sequence of *rml1*. We carried out 3'-RACE to identify 3'-end of the transcript and detected the single cDNA using separate RNA pools prepared from different tissues. Sequencing of the 1.18-kb fragment of 3' RACE product RML3' revealed a polyA signal element (AAUAAA) at position – 17 from the polyadenylation site (CA) (Fig. 1A). We finally finished *rml1_a* cDNA sequence of 1766 bp through assembling the sequences of the overlapped PCR products (Fig. 2A). The sequence was submitted to the EMBL database with accession number AJ535489.

The ORF of *rml1_a* extends 1158 bp and the weight of the predicted RML1 protein is about 42 kDa. Homology search revealed that the RML1 had 57% and 55% identity

with the GT1a/B2F from tobacco and GT-1 from *Arabidopsis*, respectively. The HLHLH DNA binding domains of these three proteins are compared in Fig. 1C. Four closely spaced helices in this region are required for DNA binding (Lam, 1995). In RML1, a glycine-rich region lies between helix3 and helix4.

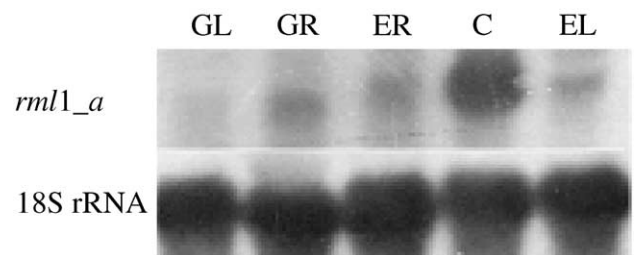


Fig. 3. Analysis of *rml1_a* RNA accumulation in different tissues of rice plants. GL, leaves of green rice plants; GR, roots of green rice plants; ER, roots of etiolated rice plants; C, calli; EL, leaves of etiolated rice plants.

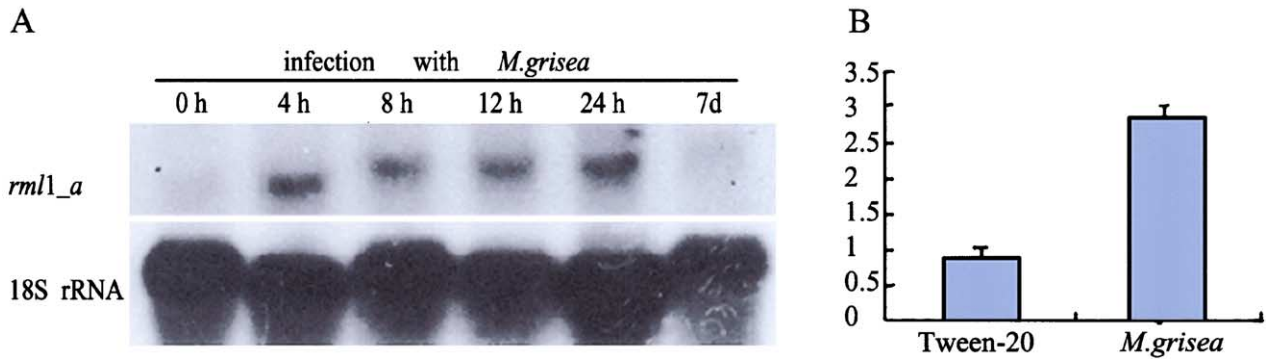


Fig. 4. *rml1_a* transcription is induced by *M. grisea*. (A) Northern blot analysis was performed on RNA extracted from seedlings incubated with the conidia of *M. grisea* at six time points (0 h, 4 h, 8 h, 12 h, 24 h, 7 days). (B) Quantitative analyses of ordinate values: Prml1 probe signal as a proportion of 18S ribosomal probe signal.

Sequence comparison between the genomic BAC clone and the corresponded cDNA clones demonstrated the exact exon–intron splicing sites (Figs. 1A and 2A). The structure of *rml1_a* is as same as that predicted by FGENESH software. The 3' UTR of *rml1_a* is confirmed with the length of 596 bp.

In addition, we also detected another alternative transcript of *rml1* gene, named as *rml1_b*, which is 1.9 kb in length (Fig. 1B). *rml1_b* has various exon–intron structures from *rml1_a* in the 5' end sequences and contains a complete intron 2 of *rml1_a* insert (Figs. 1A and 2A). The intron1 border of *rml1_b* is AU/AG (Figs. 1A and 2A). The 5' end site of *rml1_b* on the BAC sequence is different from that of *rml1_a* (Fig. 1A). Translation of all possible open reading

frames (ORF) of *rml1_b* did not yield any conceptual protein with sequence similarity to proteins in databases. The longest putative ORF from *rml1_b* is only 243 bp. So only the transcript *rml1_a* encodes a GT1-like protein.

3.2. Characterization of the promoter and the copy number of *rml1* gene

To identify the regulation of *rml1* gene expression, we cloned and sequenced a single 840-bp fragment (our submission #AJ544589) upstream of the 5' end of the sequenced BAC. Eventually, we determined 1038 bp of the nucleotide sequence upstream of the translational starting site of the *rml1* to search for *cis*-regulatory elements using

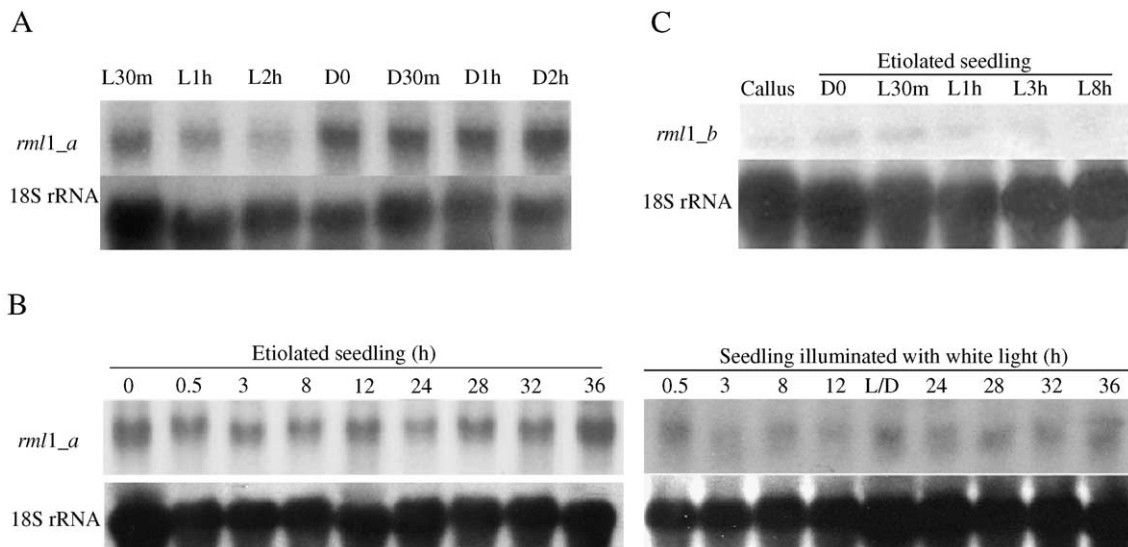


Fig. 5. Decline of *rml1* RNA levels in response to white light irradiation. The probe used for blots A and B was Prml1. (A) Seven-day-old etiolated seedlings were irradiated with white light for 30 min, 1 h, 2 h (lanes 1–3); Control etiolated seedlings were kept in the dark and harvested at the same point time (0, 30 min, 1 h, 2 h) (lanes 4–7). (B) Seven-day-old etiolated seedlings were irradiated with white light for 30 min, 3 h, 8 h, 12 h, 24 h, 28 h, 32 h, 36 h; L/D represents seedlings treated with 12 h white light then placed in the dark for 12 h before harvesting (right blot). Control etiolated seedlings were kept in the dark and harvested at the same point time (0, 30 min, 3 h, 8 h, 12 h, 24 h, 28 h, 32 h, 36 h) (left blot). (C) Northern blot analysis of *rml1_b* transcription with specific probe Prml2. Seven-day-old etiolated seedling was illuminated with white light for 0 min (lane 2), 30 min (lane 3), 1 h (lane 4), 3 h (lane 5), or 8 h (lane 6) before harvesting.

the MatInspector program against the Transfac or PLACE database (Quandt et al., 1995; Higo et al., 1999) (Fig. 2B). Regulatory elements were selected with stringent search criteria requiring 100% core factor identity and a matrix identity of more than 80% (Table 1). A TATA-like box was found between -431 and -428 in close vicinity of the putative translation start site, numbered as $+1$ (Fig. 2B). Thirteen Dof sequences may serve as binding sites for the single zinc finger transcription factors (Yanagisawa and Schmidt, 1999). A cluster of three potential elements for Opaque-2 (O_2 , transcriptional activator of Zea Mays), G-/C-box bZip and SBF-1 proteins were found (Izawa et al., 1993; Lawton et al., 1991). Other clusters of potential *cis*-element for Myb-like protein, AHBP (Arabidopsis Homeo-Box protein), RAV1 and WRKY were also observed (Solano et al., 1995; Sessa et al., 1993; Kagaya et al., 1999; Eulgem et al., 1999). In addition, the analyzed sequence contained some light related *cis*-elements such as I-box, Rebeta and Realpha elements.

Southern hybridization performed on digests of the total DNA indicated that *rml1* might be a single copy gene in the rice genome (Fig. 2C). Digestion with three different enzymes resulted in a single hybridizing fragment, with the exception of *Pst*I, where the two bands obtained reflect a *Pst*I site within the probe and total digestion at this site. In addition to RML1 and GT-2, another protein (AAK39576) with a putative trihelix motif has been predicted on chromosome10 in rice. We identified the chromosomal locations of *rml1* and *GT-2* in rice genome through searching the rice *Oryza sativa japonica* database. The results showed that these two genes are cloned in two separate BAC clones AL606587 and AL606998, respectively. Both BACs are located on chromosome 4 long arm. The BAC sequence of AL606587 containing *GT-2* is located at 83 centimorgan (cM) and the BAC sequence of AL606998 containing *rml1* is positioned at 74.2 cM. Results from database searching and Southern blot analysis confirmed that the *rml1* is a single copy gene in rice genome.

3.3. Tissue-specific expression

To identify if the *rml1_a* has tissue-specific expression, we prepared total RNA from five different tissues. The RNA gel blot revealed that the highest *rml1_a* mRNA level was detected in calli among the tested tissues. Whereas it is expressed in etiolated seedling leaves, roots of both green-grown and dark-grown seedling at lower levels. Transcription of the *rml1_a* (1.4 kb) in different tissues of rice plants is shown in Fig. 3.

3.4. Expression of *rml1* in green leaves is induced by infections with *M. grisea*

When the *rml1* gene sequence was used as a BLAST query against the EST database, most of identified ESTs were from rice leaf cDNA library infected with *M. grisea*.

Abundance of the corresponded ESTs indicated that rice blast fungus *M. grisea* might induce the expression of *rml1*. To confirm this, 4-week-old rice was inoculated with *M. grisea* and transcription of *rml1* was analyzed by Northern blot (Fig. 4A). The *rml1_a* transcript appeared 4 h after inoculation with *M. grisea*. Diseased spots appeared on leaves after spraying with *M. grisea* conidia for 7 days, but the *rml1_a* transcript was not detected in the infected leaves. Detergent alone does not increase the transcript level at the same time course (data not shown). Transcription of *rml1_a* induced by *M. grisea* is threefold higher than that of the control group (Fig. 4B). Using the Prml2 probe, we failed to detect *rml1_b* transcripts (data not shown).

3.5. *rml1* is down-regulated by light and expressed rhythmically

Sequence comparison between RML1, GT-1 and GT1a/B2F revealed striking homology. The latter two factors are

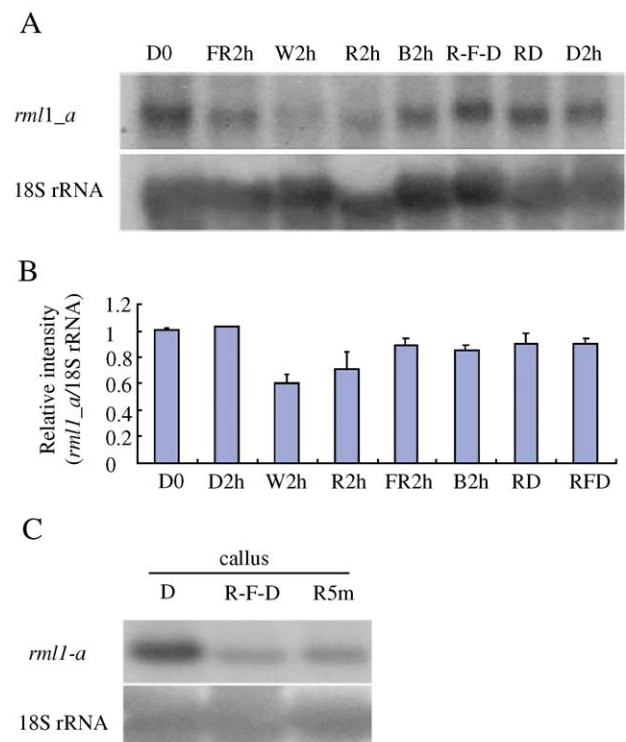


Fig. 6. The *rml1_a* expression pattern under light with Prml1 as probe. (A) Northern blot analysis of 7-day etiolated seedlings subjected to treatments of various illumination at 27 °C. Seedlings were treated with 2 h far-red light (lane 2), 2 h white light (lane 3), 2 h red light (lane 4), 2 h blue light (lane 5), 5 min pulse of red light immediately followed by 5 min far-red light then placed in the dark for 2 h (lane 6), 5 min pulse of red light then incubated in the dark for 2 h (lane 7). Lane 1 and lane 8 are control seedlings, which were kept continually in darkness. Five seedlings in each group were harvested after being illuminated. (B) Quantification of transcript levels normalized to the 18S rRNA signal. Data are expressed as mean + S.E.M. ($n \geq 3$). (C) Expression of *rml1_a* in calli. D, dark; R-F-D, 5 min red light then 5 min far-red light followed by 2 h dark; R5m, 5 min red light followed by 2 h dark.

DNA binding proteins and their cognate DNA sequences are *cis*-acting elements in light-controlled genes. This raised the interesting question of whether the rice *rml1* gene is controlled by light. After 7 days in the dark, etiolated seedlings were transferred to illumination under white light; the accumulation of *rml1_a* RNA in etiolated seedlings and irradiated seedlings were determined (Fig. 5A). The hybridization signals revealed a significant reduction in *rml1_a* RNA level in the period from 1 to 2 h following the transfer from darkness to light. We next examined the accumulation of *rml1_a* RNA in etiolated seedlings and irradiated seedlings at time-course for 36 h. The *rml1_a* level in illuminated seedlings dampens to a low, steady-level compared side-by-side with that of non-irradiated etiolated seedlings and longer periods of illumination (up to 36 h) induced a small further decrease (Fig. 5B). The *rml1_a* RNA level in seedlings first irradiated with 12 h white light then kept in darkness for 12 h, was not restored by dark (Fig. 5B). The weak expression of *rml1_b* (1.9 kb) can be detected in calli, 7-day etiolated leaves and those treated with 30 min white light (Fig. 5C).

Transcription of rice *GT-2* has also been shown to be down-regulated under white light but is unaffected by red light. To identify the phytochrome system potentially involved, we measured *rml1_a* transcription levels in etiolated seedlings and seedlings irradiated with red, far-red or blue light. Although the transcription level did not change visibly during 30 min illumination (data not shown), all light treatments elicited a reduction in *rml1_a* RNA levels after

2 h irradiation (Fig. 6A,B). Transcription of *rml1_a* in etiolated seedling dropped to 60% of the dark level under white light and to about 72% of the dark level under red light (Fig. 6B). A significant difference in the regulation of *rml1_a* levels by red light was apparent between calli and etiolated seedlings. In calli, a 5-min pulse of red light dramatically repressed the *rml1_a* expression (Fig. 6C).

In the untreated control in *M. grisea* infection experiment, we noticed the weak expression in *rml1_a* mRNA at 16 o'clock and 20 o'clock (data not shown), suggesting that *rml1* may be under circadian regulation. The rhythmic pattern of *rml1_a* RNA was not detected in 7-day-old etiolated seedling (Fig. 5B). When its temporal expression pattern was analyzed with 17-day-old green seedling, we observed a rhythm of *rml1_a* abundance over 2.5 days in LD cycles. Fig. 7A,B shows that in LD 13:11 (lights on from 6 o'clock in morning to 19 o'clock in evening), the *rml1_a* is expressed rhythmically with highest expression at approximately 20 o'clock (evening) and the lowest expression occurred at 8 o'clock (morning) in a time-course experiment. To investigate if the rhythmic cycling of *rml1_a* RNA levels was under the circadian clock control, these seedlings grown in L/D were transferred to continuous light (LL) or continuous dark (DD) for 24 h and then assayed for *rml1_a* expression over a 48-h period. *rml1_a* transcript has no detectable rhythms in LL or DD conditions (Fig. 7C). The expression level of *rml1_a* in DD increased compared with the *rml1_a* RNA level in LL over the first 24-h period (Fig. 7C).

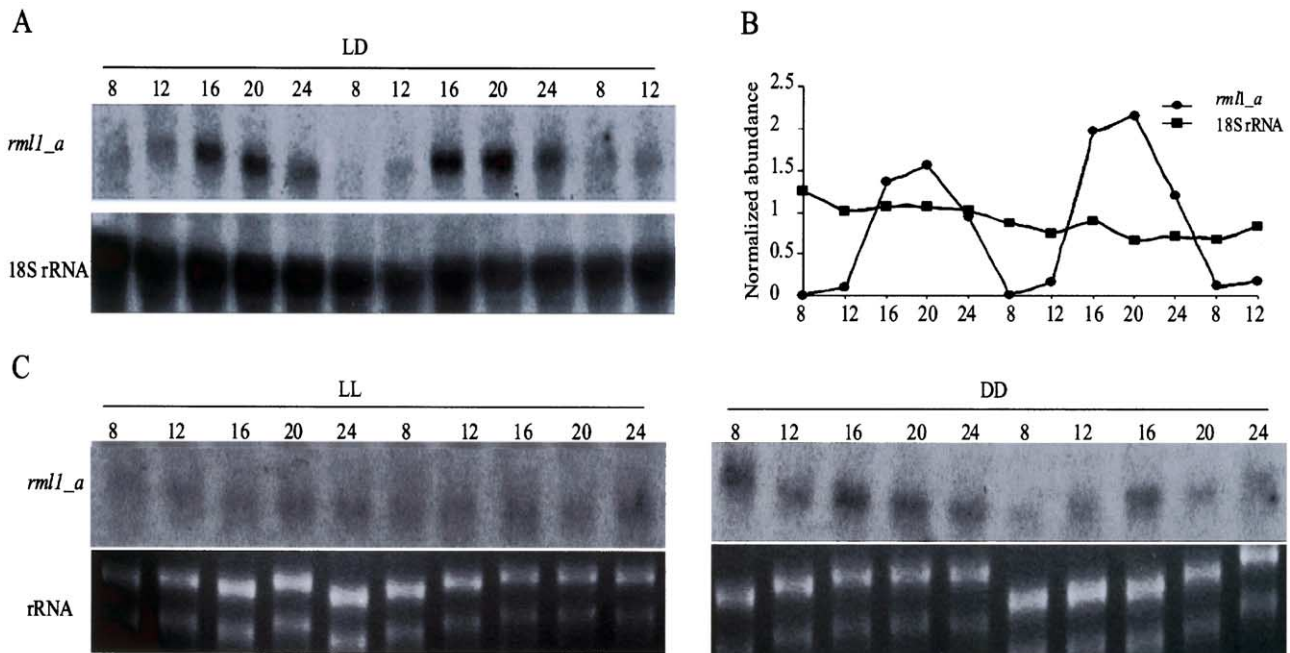


Fig. 7. The daily rhythms of *rml1_a* RNA abundance. (A) Rice germinated seeds were first grown in the dark for 3 days and then entrained to LD cycle for 2 weeks before being harvested RNA from leaves at the indicated hours (o'clock). (B) Quantification of the data shown in (A). (C) The 17-day seedlings grown in LD cycle were transferred to continuous light (LL) or continuous dark (DD) for 24 h before being extracted RNA at the indicated hours (o'clock) over 48 h. The bottom panels show 28S and 18S rRNA bands in the ethidium bromide-stained gel.

4. Discussion

We have cloned a new gene, *rml1*, which is a single copy gene based on the results of the Southern blot analysis and rice genome database search. Chromosomal locations of the *rml1* and *GT-2* genes were identified at 74.2 and 83 cM on the long arm of chromosome 4, respectively. The transcription of *rml1* was identified by RNA gel analysis. In our studies, we found that the constructed *rml1_a* cDNA is longer (1766 bp) than RNA analyzed by Northern blot analysis (about 1.4 kb). Our experiments and other evidence supposed that the length discrepancy between constructed cDNA and RNA detected on the formaldehyde gel may be resulted from the secondary structure of *rml1_a* in gel (Dawson et al., 1996).

The *rml1* expression in green leaves was up-regulated by the rice blast fungus *M. grisea* and this induction is in the early stages of *M. grisea* infection (Fig. 4). A W-box (TTGACT/A) has been found to be important in the plant defense response to pathogens (Rushton et al., 2002). The W-box binding protein WRKY is suggested to regulate the early defence-response genes (Eulgem et al., 1999). The presence of W-box in *rml1* promoter may play a role in the *M. grisea* regulation of *rml1* expression. The conidia of *M. grisea* are germinated after 1 ~ 2 h and form hook after 6–12 h (by communications with Dr. Eckhard Thines). So it is interesting to understand the interaction between rice and conidia and the mechanism of *rml1_a* expression induced by conidia.

Expression of *rml1_a* in 17-day green seedling grown in L/D then transferred to continuous light or continuous dark shows no circadian expression (Fig. 7C). These data suggest that LD cycling is essential in maintaining rhythm control of *rml1*. The importance of LD cycles in entraining the rhythm expression of *rml1_a* was confirmed when *rml1_a* mRNA did not show rhythm regulation in 7-day-old etiolated seedling that germinated and grown in continuous dark (Fig. 5B). Like *rml1_a*, the rhythm regulation of cysteine proteinase gene *SmCP* is dependent on LD cycle (Xu et al., 2003). The circadian expression of some genes are independent on LD cycle because seedling grown in LD and transferred in continuous light showed circadian regulation, suggesting that clock may act to rhythmic regulation of these gene expression (Fowler et al., 1999). Two G-box cores like elements, in which the first position changes from C to T, were found at –857 and –428 positions (as shown in Table 1) in *rml1* gene promoter. G-box core sequences (CACGTG) functions as clock-regulated motifs in tobacco circadian clock gene promoter (Xu and Johnson, 2001). Dissection of interactive *cis*-elements and transcription factors may uncover the mechanisms of variations in circadian regulation of genes expression.

In etiolated seedlings, a significant reduction in transcript levels was detected after continuous irradiation with red light. The results suggest that at least one light-stable phytochrome may mediate light regulation of *rml1_a* in

etiolated seedlings. In contrast, *rml1_a* expression in calli shows obvious red pulse treatment phenomena, in which the RNA level declines dramatically under a 5 min pulse of red light (Fig. 6C). Since *rml1_a* transcript levels changed differently in response to various irradiation conditions, we propose that the repression of *rml1_a* mRNA accumulation in calli or etiolated seedlings is regulated via phytochromes in different ways. Phototransduction pathways are possibly controlled by interacting photoreceptors (Más et al., 2000). The expression level of *rml1_a* RNA in etiolated seedling was regulated by light of various wavelengths with different effectiveness (Fig. 6A,B). These results indicated that other photoreceptors in addition to the red light receptor phyB are also involved in the regulation of *rml1_a* RNA levels in etiolated seedlings.

In summary, we characterized the expression pattern of a novel GT1-like factor gene *rml1* in rice. RML1 is unique among single trihelix factors in that its RNA level is down-regulated by light in a phytochrome-related manner. Besides, the expression of *rml1* is induced by *M. grisea* infection and it was expressed rhythmically.

Acknowledgements

We are grateful to Zuhua He and Hongquan Yang (Shanghai Institute of Plant Physiology and Ecology) for assistance on pathogenesis analysis of the rice infection by blast *M. grisea* conidia and for providing the light instruments. We thank Huatai Xu for statistical analysis of figures and useful discussions. We are also grateful to the referees and the editor for their helpful comments and suggestions. This work was supported by grants from the Ministry of Science and Technology of the People's Republic of China, Chinese Academy of Sciences, and the Shanghai Municipal Commission of Science and Technology.

References

- Buchel, A.S., Brederode, F.T., Bol, J.F., Linthorst, H.J., 1999. Mutation of GT-1 binding sites in the Pr-1A promoter influences the level of inducible gene expression in vivo. *Plant Mol. Biol.* 40, 387–396.
- Dawson, M.T., Powell, R., Gannon, F., 1996. *Gene Technology*. Bios Scientific Publishers, Information Press, Oxford, UK, pp. 87–88.
- Dehesh, K., Bruce, W.B., Quail, P.H., 1990. A *trans*-acting factor that binds to a GT-motif in a phytochrome gene promoter. *Science* 250, 1397–1399.
- Eulgem, T., Rushton, P.J., Schmelzer, E., Hahlbrock, K., Somssich, I.E., 1999. Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. *EMBO J.* 18, 4689–4699.
- Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G., Putterill, J., 1999. *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. *EMBO J.* 18, 4679–4688.
- Gilmartin, P.M., Memelink, J., Hiratsuka, K., Kay, S.A., Chua, N.H., 1992.

- Characterization of a gene encoding a DNA binding protein with specificity for a light responsive element. *Plant Cell* 4, 839–849.
- Higo, K., Ugawa, Y., Iwamoto, M., Korenaga, T., 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res.* 27, 297–300.
- Hiratsuka, K., Wu, X., Fukuzawa, H., Chua, N.H., 1994. Molecular dissection of GT-1 from *Arabidopsis*. *Plant Cell* 6, 1805–1813.
- Izawa, T., Foster, R., Chua, N.H., 1993. Plant bZIP protein DNA binding specificity. *J. Mol. Biol.* 230, 1131–1144.
- Kagaya, Y., Ohmiya, K., Hattori, T., 1999. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.* 27, 470–478.
- Kuhn, R.M., Caspar, T., Dehesh, K., Quail, P.H., 1993. DNA binding factor GT-2 from *Arabidopsis*. *Plant Mol. Biol.* 23, 337–348.
- Lam, E., 1995. Domain analysis of the plant DNA-binding protein GT1a: requirement of four putative α -helices for DNA binding and identification of a novel oligomerization region. *Mol. Cell. Biol.* 15, 1014–1020.
- Lam, E., Chua, N.H., 1990. GT-1 binding site confers light responsive expression in transgenic tobacco. *Science* 248, 471–474.
- Lawton, M.A., Dean, S.M., Dron, M., Kooter, J.M., Kragh, D.M., Harrison, L., Yu, L., Tanguay, L., Dixon, R.A., Lamb, C.J., 1991. Silencer region of a chalcone synthase promoter contains multiple binding sites for a factor, SBF-1, closely related to GT-1. *Plant Mol. Biol.* 16, 235–249.
- Le Gourrierec, J., Li, Y.F., Zhou, D.X., 1999. Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA–TBP–TATA complex. *Plant J.* 18, 663–668.
- Más, P., Devlin, P.F., Panda, S., Kay, S.A., 2000. Functional interaction of phytochrome B and cryptochrome 2. *Nature* 408, 207–211.
- Ni, M., Dehesh, K., Tepperman, J.M., Quail, P.H., 1996. GT-2: in vivo transcriptional activation activity and definition of novel twin DNA binding domains with reciprocal target sequence selectivity. *Plant Cell* 8, 1041–1059.
- O’Grady, K., Goekjian, V.H., Nairn, C.J., Nagao, R.T., Key, J.L., 2001. The transcript abundance of GmGT-2, a new member of the GT-2 family of transcription factors from soybean, is down-regulated by light in a phytochrome-dependent manner. *Plant Mol. Biol.* 47, 367–378.
- Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T., 1995. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* 11, 4878–4884.
- Rushton, P.J., Reinstadler, A., Lipka, V., Lippok, B., Somssich, I.E., 2002. Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen- and wound-induced signaling. *Plant Cell* 14, 749–762.
- Sessa, G., Morelli, G., Ruberti, I., 1993. The Athb-1 and -2 HD-zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO J.* 12, 3507–3517.
- Solano, R., Nieto, C., Avila, J., Canas, L., Diaz, I., Paz-Ares, J., 1995. Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*. *EMBO J.* 14, 1773–1784.
- Villain, P., Mache, R., Zhou, D.X., 1996. The mechanism of GT element-mediated cell type-specific transcriptional control. *J. Biol. Chem.* 271, 32593–32598.
- Xu, Y., Johnson, C.H., 2001. A clock- and light-regulated gene that links the circadian oscillator to *LHCB* gene expression. *Plant Cell* 13, 1411–1425.
- Xu, Z.F., Chye, M.L., Li, H.Y., Xu, F.X., Yao, K.M., 2003. G-box binding coincides with increased *Solanum melongena* cysteine proteinase expression in senescent fruits and circadian-regulated leaves. *Plant Mol. Biol.* 51, 9–19.
- Yanagisawa, S., Schmidt, R.J., 1999. Diversity and similarity among recognition sequences of Dof transcription factors. *Plant J.* 17, 209–214.
- Zhao, Q., Zhang, Y., Cheng, Z., et al., 2002. A fine physical map of the rice chromosome 4. *Genome Res.* 12, 817–823.
- Zhou, D.X., 1999. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends Plant Sci.* 4, 210–214.