Kasetsart J. (Nat. Sci.) 43 : 164 - 177 (2009)

In silico Promoter Analysis of Photoperiod-Responsive Genes Identified by DNA Microarray in Rice (*Oryza sativa* L.)

Chareerat Mongkolsiriwatana[1,2], Pradit Pongtongkam[1] and Surin Peyachoknagul[1]*

ABSTRACT

The bioinformatics tools, *in silico* analysis, were used to analyze the promoter regions of photoperiod-responsive genes in rice obtained from microarray analysis. The results revealed that the common feature of a photoperiod-responsive promoter is the combination of various light-responsive elements, hormone-responsive elements and stress-responsive elements. This indicated that the photoperiod response is controlled by light in coordination with hormones and stresses. The coordinated motifs, integrating hormones and stresses to photoperiod responses were identified. The GARE motif and G-box are the coordinated motifs integrating gibberellins with photoperiod while MBS and G-box are the coordinated motifs integrating ethylene or abscisic acid and stresses with photoperiod. Moreover, the specific organization of *cis*-regulatory elements was used to identify phytochrome A (phyA)-regulated genes. The data showed that phyA is involved in the expression of flowering-time genes either by the activation of floral inducers or the suppression of floral repressors. In addition, two novel *cis*-regulatory elements which are specific to day-length were identified. The novel A-rich element is specific to long daylight and is involved in the regulation of phyA and circadian rhythm to inhibit flowering whereas the novel GC element is specific to short daylight and involved in the regulation of gibberellin signaling to promote flowering. In conclusion, photoperiodic flowering in rice is controlled, partially, by phyA in coordination with A-rich and GC elements.

Key words: *in silico* promoter, photoperiod responsive promoter, over-representative motif, *cis*-regulatory element, rice

INTRODUCTION

In plants, a large number of transcription factors are known to control the expression of target genes in various signal transduction cascades (Venter and Botha, 2004). Transcription-factor binding sites (TFBs or *cis*-regulatory elements), on the other hand, determine the specific timing and location of transcriptional activity. They are primarily located in the long non-coding sequence upstream of a gene (Chaboute *et al*., 2002). These regulatory motifs organized into distinct *cis*-regulatory modules are required for a specific expression pattern (Babu *et al*., 2004). Thus, the identification of regulatory motifs and their organization is an important step to improve understanding of gene expression and regulation.

1 Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.

2 Department of Genetics, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand.

* Corresponding author, e-mail: fscisrp@ku.ac.th

Received date : 02/09/08 Accepted date : 29/10/08

Many databases have been developed for the prediction of *cis*-regulatory sites in a non-coding DNA sequence such as PlantCARE (Lescot *et al.*, 2002) and PLACE (Higo *et al.*, 1999). The candidate *cis*-regulatory elements in the promoter regions can be identified by searching against the known elements in the databases. In addition, novel *cis*-regulatory elements could also be detected with no prior knowledge of transcription-factor binding sites using the over-representative sequence of the promoters of co-expressed genes (Helden, 2003). Therefore, the use of *in silico* promoter analysis through bioinformatics as a tool has become more attractive and feasible when combined with DNA microarray analysis.

In a previous experiment, the GeneChip DNA microarray technique was used to investigate the global expression of photoperiod-responsive genes and to identify the photoperiodic-flowering pathway in rice (*Oryza sativa* L.) cultivar KDML 105 (Mongkolsiriwatana, 2008). However, the regulatory mechanism of the transcriptional control remained to be characterized. In this study, therefore, the bioinformatics tools were used to analyze the promoter of these genes. All known *cis*-regulatory elements that occurred in the promoter regions were investigated and also the coordinated motifs integrating plant hormones and stresses with photoperiod responses were identified. Furthermore, the specific organization of *cis*-regulatory elements was used to detect phytochrome A (phy A)-regulated genes from the photoperiod-responsive pathway and two novel *cis*-regulatory elements specific to the day length in rice were elucidated.

MATERIALS AND METHODS

Promoter sequence data sets

The photoperiod-responsive genes in rice identified and used in this study were derived from the data of Mongkolsiriwatana (2008), and their orthologous genes in *Arabidopsis* were identified using rice-nucleotide sequences searched against *Arabidopsis* genes in the TAIR8 database (http://www.arabidopsis.org/wublast/index2.jsp). Based on the gene location from the Rice TIGR release 5 or TAIR8 database, the 1.5 kb sequence upstream from the start site was extracted as a putative promoter region. The data sets consisted of the promoter sequences of rice from induced (61 sequences) and suppressed (81 sequences) genes under SD light, which were named as SD and LD promoters, respectively, as well as those of the SD and LD orthologous genes in *Arabidopsis*.

Search for known *cis*-regulatory elements

The putative promoter sequences of photoperiod-responsive genes were compared against known *cis*-regulatory elements in the collection of the PlantCARE database using the Search for Care program (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The *cis*-regulatory elements were listed and the occurrence number on each promoter was recorded. The percentage of these specific elements relative to the total number in each data set was calculated.

Analysis of over-representative motifs from co-expressed genes

Each promoter data set of co-expressed genes was used to detect the over-representative sequence using an oligo-analysis program (Helden *et al.*, 1998; Helden, 2003). The oligomer counting mode was set at 10 oligomer lengths, to prevent overlapping matches and counting on a single strand. The consensus sequence output from the oligo-analysis was converted to a logo picture by the Gibbs program (Helden, 2003). The over-representative sequences were analyzed for any positional bias using a position-analysis program (Helden, 2003). These programs are available in the Regulatory Sequence Analysis Tool (RSAT) package (http://rsat.scmbb.ulb.ac.be/rs at/).

RESULTS AND DISCUSSION

Common features of photoperiod-responsive promoters

To characterize the general features of the promoter regions of the photoperiod-responsive genes, the 1.5-kb sequences upstream from the start site of the induced and suppressed genes were used to search against known *cis*-regulatory elements in the PlantCARE database using the Search for Care program (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The specific elements presented in the promoter regions of the coding strands were counted and are listed in Table 1. The data revealed that these promoter regions contained several *cis*-regulatory elements; 839 in SD promoters and 1,231 in LD promoters. These elements were classified into three groups according to their responsive functions: a hormone-responsive element (HRE); a light-responsive element (LRE); and a stress-responsive element (SRE). The HREs were composed of five subgroups: abscisic acid (ABA)-responsive elements, gibberellins (GA)-responsive elements, auxin-responsive elements, jasmonic acid (JA)-responsive elements and ethylene-responsive elements as shown in Table 1. The data also indicated that the transcription of photoperiod-responsive genes was specifically influenced not only by light, but also by various hormones and stresses, which agreed with a previous report (Mongkolsiriwatana, 2008). This showed the response of photoperiod crosstalk with stresses and hormones.

To compare the occurrence of elements in a subset of the genome, the percentage of individual elements in each set of promoters was calculated and is depicted in Figure 1. As expected, LREs which commonly occurred in light-regulatory promoters (Jiao *et al*., 2007) were enriched in the promoters of photoperiod-responsive genes. Thus, these data showed that

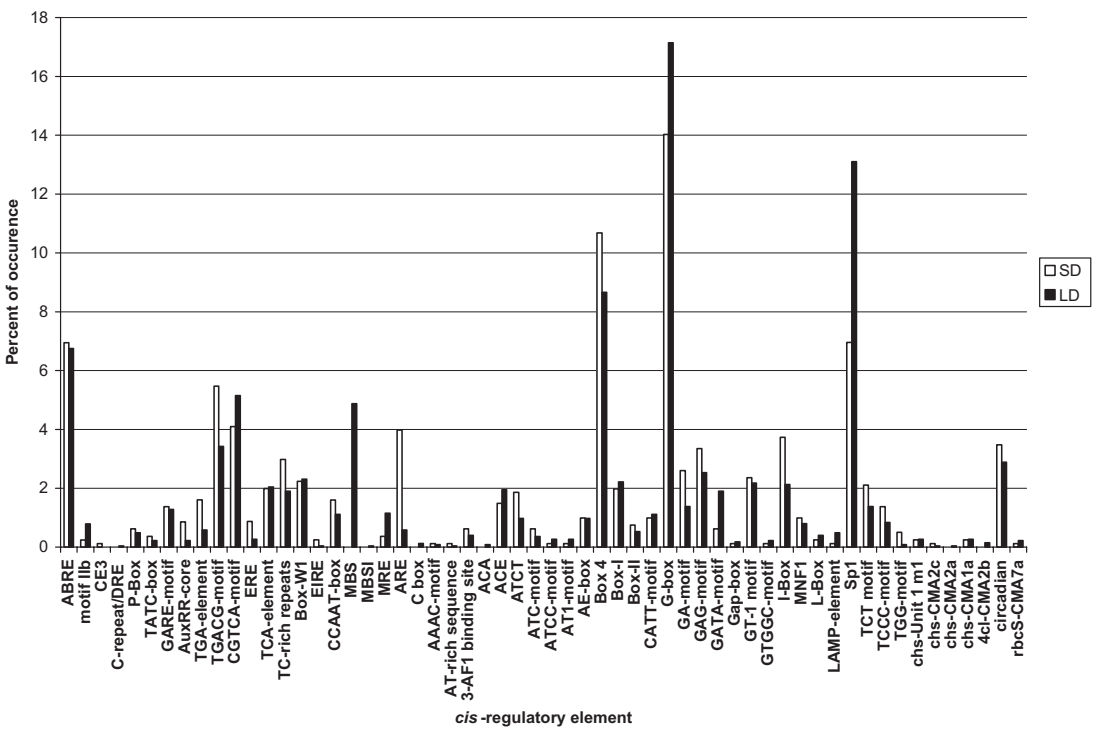

Figure 1 The distribution occurrence of *cis*-regulatory elements in the promoter of photoperiod responsive genes.

Table 1 The collection of *cis*-regulatory elements in the promoter of photoperiod-responsive genes.

Type of motif	Sequence	Function	Total number of occurrence	
			SD	LD
Abscisic acid-responsive element				
ABRE	TACGTG	cis-acting element involved in the abscisic acid responsiveness	56	78
motif IIb	CCGCCGCGCT	Abscisic acid responsive element	2	8
CE3		ABA and VP1 responsive	1	0
Gibberellins-responsive element				
P-Box	CCTTTTG	Gibberellin-responsive element	5	6
TATC-box	TATCCCA	Cis-acting element involved in gibberellin responsiveness	3	1
GARE-motif	AAACAGA	Gibberellin-responsive element	11	18
Auxin-responsive element				
AuxRR-core	GGTCCAT	cis-acting regulatory element involved in auxin responsiveness	7	1
TGA-element	AACGAC	Auxin-responsive element	13	3
Jasmonic acid-responsive element				
TGACG-motif	TGACG	cis-acting regulatory element involved in the MeJA-responsiveness	44	43
CGTCA-motif	CGTCA	cis-acting regulatory element involved in the MeJA-responsiveness	33	71
Ethylene-responsive element				
ERE	ATTTCAAA	Ethylene-Responsive Element	7	4
TCA-element	CCATCTTTTT	cis-acting element involved in salicylic acid responsiveness	16	26
TC-rich repeats	ATTCTCTAAC	cis-acting element involved in defense and stress responsiveness	24	22
Box-W1	TTGACC	fungal elicitor responsive element	18	35
EIRE	TTCGACC	Elicitor-responsive element	2	1
Stress-responsive element				
CCAAT-box	CAACGG	MYBHv1 binding site	13	14
MBS	TAACTG	MYB binding site involved in drought inducibility	34	62
ARE	TGGTTT	Cis-acting regulatory element essential for the anaerobic induction	32	6
Light-responsive element				
MRE	AACCTAA	MYB binding site involved in light responsiveness	3	11
C box	CTGACGTCAG	Cis-acting regulatory element involved in light responsiveness	0	3
AAAC-motif	CAATCAAAACCT	light responsive element	1	1
AT-rich sequence	ATTAATTTTACA	part of a light responsive module	1	1
3-AF1 binding site	TAAGAGAGGAA	light responsive element	5	6
ACE	CTAACGTATT	Cis-acting element involved in light responsiveness	12	24
ATCT	AATCTAATCT	part of a conserved DNA module involved in light responsiveness	15	14
ATC-motif	AGCTATCCA	part of a conserved DNA module involved in light responsiveness	5	4
ATCC-motif	CAATCCTC	part of a conserved DNA module involved in light responsiveness	1	3
AT1-motif	AATTATTTTTTATT	part of a light responsive module	1	4
AE-box	AGAAACAT	part of a module for light response	8	10
Box 4	ATTAAT	part of a conserved DNA module involved in light responsiveness	86	113
Box-I	TTTCAAA	light responsive element	16	23
Box-II	TGGTAATAA	part of a light responsive element	6	9
CATT-motif	GCATTC	part of a light responsive element	8	11
G-box	CACGTA	Cis-acting regulatory element involved in light responsiveness	113	208
GA-motif	AAAGATGA	part of a light responsive element	21	14
GAG-motif	AGAGATG	part of a light responsive element	27	36
GATA-motif	GATAGGA	part of a light responsive element	5	19
Gap-box	CAAATGAA(A/G)A	part of a light responsive element	1	1
GT-1 motif	GGTTAAT	light responsive element	19	24
GTGGC-motif	GATTCTGTGGC	part of a light responsive element	1	2
I-Box	GATATGG	part of a light responsive element	30	22
MNF1	GTGCCC(A/T)(A/T)	light responsive element	8	11
L-Box	TCTCACCAACC	part of a light responsive element	2	4
LAMP-element	CCAAAACCA	part of a light responsive element	1	7
□Sp1	CC(G/A)CCC	Light responsive element	56	176
TCT motif	TCTTAC	part of a light responsive element	17	13
TCCC-motif	TCTCCCT	part of a light responsive element	11	13
TGG-motif	GGTTGCCA	part of a light responsive element	4	2
chs-Unit 1 m1	ACCTAACCTCC	part of a light responsive element	2	1
chs-CMA2c	ATATACGTGAAGG	part of a light responsive element	1	1
chs-CMA2a	TCACTTGA	part of a light responsive element	0	1
chs-CMA1a	TTACTTAA	part of a light responsive element	2	3
4cl-CMA2b	TCTCACCAACC	Light responsive element	0	1
rbcS-CMA7a	GGCGATAAGG	part of a light responsive element	1	1
□Circadian	CAANNNNATC	cis-acting regulatory element involved in circadian control	28	35
		Total regulatory elements	**839**	**1231**

the expressions of photoperiod-responsive genes were mainly controlled by LREs in coordination with plant HREs and SREs. This agreed with the obtained microarray data (Mongkolsiriwatana, 2008), in which the expression of some hormone-signaling and stress-responsive genes was altered by SD light.

Among LREs themselves, G-box, Sp1 and Box-4 were predominantly found in both SD and LD promoters, suggesting that these three elements were the common *cis*-regulatory elements of photoperiod-responsive genes. Interestingly, G-box, a target site of phytochrome interacting factors (PIFs), which is required for phytochrome-regulated transcription in photoperiod response (Menkens *et al*., 1995), as well as for stress and defense responses (Aris *et al*., 1993) was the highest enriched one. Thus, this observation supported the hypothesis that the expression of photoperiod-responsive genes was mainly controlled by phytochromes. In addition, circadian elements were found in both SD and LD promoters. These data indicated that some photoperiod-responsive genes are controlled by the circadian clock as also reported by Young and Kay (2001).

Obviously, no single element of LREs was found in any photoperiod-regulated promoters, suggesting that a combination of the light-regulatory network is required for the photoperiod response. The specific combination of LREs for a photoperiod response could not be identified in this experiment, because of its complexity. However, this study is the first report to show the characteristics of the photoperiod-responsive promoter.

Identification of coordinated motifs integrating hormones and stresses to photoperiod responses

To further understand the coordinated regulation between photoperiod and hormones, the promoters of hormone-signaling genes obtained from microarray data were analyzed. These genes were *SLEEPY 1* (*SLY 1*) in GA signaling with up-regulation, *XRN 4* in ethylene signaling and *Viviparous 14* (*Vip 14*) in ABA signaling with down-regulations. All *cis*-regulatory elements that occurred on the promoter region of these genes were mapped as shown in Figure 2. Several LREs were identified in each of the three promoters but with different combinations. Interestingly, the GA-responsive element (GARE motif) was found in

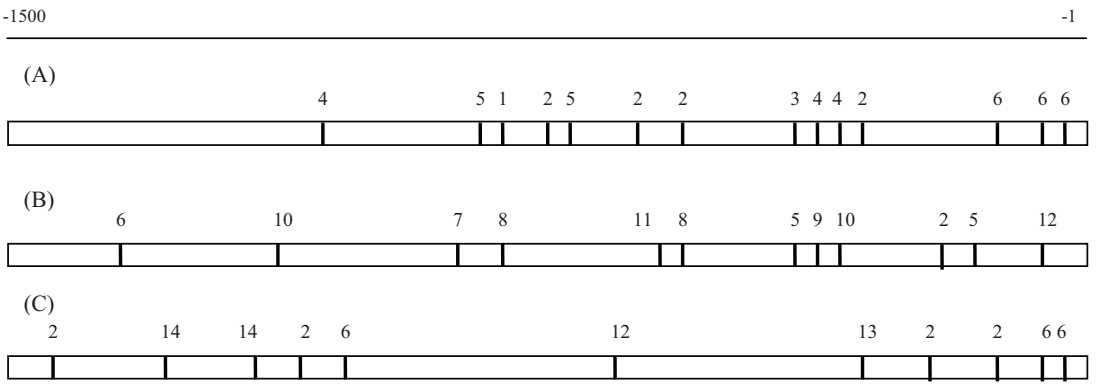

Figure 2 Feature map of *cis*-regulatory elements on the promoter region of *SLY 1* (A), *XRN 4* (B) and *Vip 14* (C). The Arabic numbers 1 to 14 indicate *cis*-regulatory elements; 1= ACE element, 2= G-box, 3 = GARE, 4 = LTR, 5 = MRE, 6= Sp1, 7 = AT1 motif, 8 = Box 4, 9 = CCGTCC-box, 10 = GAG motif, 11 = GT-1 motif, 12 = MBS, 13 = AE-box and 14 = GATA. The function of the *cis*-regulatory elements is shown in Table 1.

the promoter of the *SLY 1* gene, supporting that view that the transcription of *SLY 1* was induced by GA in coordination with light responses. The expressed product of the *SLY 1* gene is the F-box subunit of SCF E3 ubiquitin ligase that activates GA signaling by promoting the proteolysis of a GA repressor (DELLA protein) (Lechner *et al*., 2006). The expression of *SLY 1* is known to be controlled by the GA level (Hirano *et al*., 2008). Thus, this evidence suggested that the GA level was increased by SD light, leading to the activation of *SLY 1* transcription via the GARE motif on its promoter. This hypothesis was supported by the work of Kulikowska-Gulewska *et al*. (2002), which demonstrated that the GA level was increased by SD light in *Pharbitis nil* (an SD plant). Furthermore, G-Box which is a recognition site for phytochrome-interacting factors (PIFs) was found in the *SLY 1* promoter. This showed that phytochromes are also involved in the regulation of the *SLY 1* expression in the GA-signaling response. It has also been reported that phytochromes were mainly involved in the regulation of the photoperiod response in rice (Izawa *et al*., 2007). Bringing all this information together, it could be concluded that the expression of *SLY 1* is an integrator, linking the GA hormone to the photoperiod via the specific organization of G-box and GARE on its promoter. Thus, G-box and GARE are the coordinated motifs which integrate the GA hormone to photoperiod responses.

In addition, the MBS motif, which is the target site of the myb-transcription factor, responded to abiotic stresses and with G-box was found in the promoters of *XRN 4* and *Vip 14* (Figure 2). This showed that ethylene and ABA signaling are controlled by phytochrome in the photoperiod and stress responses, which was in agreement with previous reports, that demonstrated both ethylene and ABA-signaling were involved in abiotic stresses (Rabbani *et al*., 2003; Achard *et al*., 2006; Zhou *et al*., 2007; Huang *et al*., 2008) and in photoperiod responses (Welling *et al*., 1997; Kesy *et al*., 2008). It suggested that the expression of *XRN 4* and *Vip 14* are integrators linking stress-responsive genes to photoperiodic regulation. Thus, MBS and G-box are coordinated motifs which integrate ethylene or ABA hormones and stresses to photoperiod responses.

Identification of phytochrome A-regulated genes using specific motifs in their promoter regions

Phytochrome A (phyA)-signaling genes, i.e., *PIF 3*, *FAR 1*, *COP 1* were found in photoperiod-responsive genes, showing that phyA must be involved in the photoperiod response in rice. Thus, it is interesting to further identify the phyA-regulated genes to elucidate the influence of phyA in photoperiod responsive pathway. The promoters of phyA-regulated genes are known to contain the G-box and GT-1 motifs and/or the GATA motif, while the G-box itself was found to be located upstream of the GT-1 motif and/or GATA motif (Hudson and Quail, 2003; Jiao *et al*., 2007). Thus, these criteria were used to select the phyA-regulated genes from the members of the SD induced and SD-suppressed genes. These genes are listed in Table 2 and Table 3 respectively. The data indicated that the function of phyA was not only to activate, but also to suppress the expression of genes as suggested by Tepperman *et al*. (2001). Interestingly, both the flowering time-regulated genes, i.e, floral inducers (*Hd3a*, *MADS box transcription factor I* and *MADS box transcription factor 14*) in the SD-induced genes and floral repressors (*AP 2*, *FAR 1* and *Vip 14*) in the SD-suppressed genes were found. This result supported the hypothesis that phyA regulates flowering time in response to photoperiod either by the direct or indirect activation of floral inducers or the suppression of floral repressors in rice. This finding is consistent with the notion that phyA plays a promotional role in flowering as seen from the early flowering of over-expressed *phyA*, but

Table 2 Putative phytochrome A–regulated genes induced by SD light.

RepresentativeID	TIGR locus	Description
AK101991	LOC_Os12g09700	Jasmonate-induced protein
AK069331	**LOC_Os03g54160**	**MADS-box transcription factor 14**
AK067894	LOC_Os10g33290	NPGR2, putative, expressed
AK068816	LOC_Os08g10310	SHR5-receptor-like kinase
L34271	**LOC_Os03g11614**	**MADS-box transcription factor 1**
AK099709	LOC_Os03g61160	Expressed protein
U30479	LOC_Os05g19570	Alpha-expansin 1 precursor
AK107926	LOC_Os01g28450	Pathogenesis-related protein PRB1-2
AK102086	LOC_Os10g40440	Cortical cell-delineating protein precursor,
AK065631	LOC_Os11g45990	Protein binding protein, putative
AB052943	**LOC_Os06g06320**	**OsFTL2 – Rice FT-Like2 (Hd3a)**
AK107044	LOC_Os04g11660	F-box protein interaction domain
AB017914	LOC_Os11g42090	Leucine Rich Repeat family protein
BI798695	LOC_Os03g55734	Prolamin, putative, expressed
AK100128	LOC_Os04g03796	peptidase/ subtilase, putative, expressed
AK059202	LOC_Os11g10590	Expressed protein
9629.m06682	LOC_Os01g67280	Hypothetical protein
9630.m05489	LOC_Os03g23040	Hypothetical protein
9639.m03143	LOC_Os11g34990	Hypothetical protein
9640.m02992	LOC_Os12g30940	F-box domain containing protein
AK067428	LOC_Os02g48340	RNA-binding region-containing protein
NM_190345	LOC_Os01g60510	Hypothetical protein
9631.m02338	LOC_Os03g23940	Expressed protein
9631.m04608	LOC_Os01g07180	Hypothetical protein
9640.m00101	LOC_Os12g01990	Expressed protein

Note: Floral inducing genes are highlighted by **bold font.**

the late flowering of mutated *phyA* in *Arabidopsis* (Johnson *et al*., 1994; Bagnall *et al*., 1995). Taken all together, this study proposed that phyA is involved in the regulation of photoperiodic flowering by promoting floral inducers while inhibiting floral repressors via the specific organization of *cis*-regulatory elements on the promoter of the target genes.

Identification of novel *cis*-regulatory elements from the promoter of photoperiod responsive genes

To identify the novel *cis*-regulatory elements that are a specific response to day-length in rice, the promoter region of co-expressed genes in each set was used to identify a common over-representative motif using oligo-analysis and subsequently converted to a logo picture by the Gibbs programs. The results showed that the GCCGGCGCCC and AAAAAAAAAA elements were over-represented in the promoter of SD and LD genes, respectively (Figure 3). The GCCGGCGCCC combination was assigned as the GC element while AAAAAAAAAA was assigned as the A-rich element. Whether or not these over-represented sequences were positionally biased in the promoter of the SD or LD genes, the distribution of each element in the promoter of

Table 3 Putative phytochrome A–regulated genes suppressed by SD light.

RepresentativeID	TIGR locus	Description
9632.m04031	LOC_Os04g42260	Protein phosphatase 2C isoform gamma
9635.m03511	LOC_Os07g35540	Protein kinase, putative, expressed
AK058773	LOC_Os02g08440	OsWRKY71 – Superfamily of rice TFs
AK060639	LOC_Os10g36500	Pectinesterase inhibitor domain
AK063042	LOC_Os03g08520	Expressed protein
AK064287	LOC_Os02g54590	Serine threonine kinase, putative
AK067836	LOC_Os09g31130	Tonoplast dicarboxylate transporter
AK069654	LOC_Os02g18064	expressed protein
AK101750	LOC_Os01g43740	cytochrome P450 72A1, putative,
AK109161	LOC_Os03g13740	U-box domain containing protein,
AK110739	LOC_Os10g36520	expressed protein
AK111335	LOC_Os02g50110	expressed protein
AK111571	LOC_Os01g64360	Myb-like DNA-binding domain
AK119780	LOC_Os07g05940	viviparous-14, neoxanthin cleavage enzyme
AY345234	LOC_Os01g73770	Transcription factor RCBF2, AP 2 protein
BP184627	LOC_Os01g19050	far-red impared
CB621890	LOC_Os02g52390	neoxanthin cleavage enzyme
NM_188146	LOC_Os01g17396	expressed protein
AK108588	**LOC_Os06g49550**	**far-red impaired responsive protein**

Note: Floral repressing genes are highlighted by **bold font**

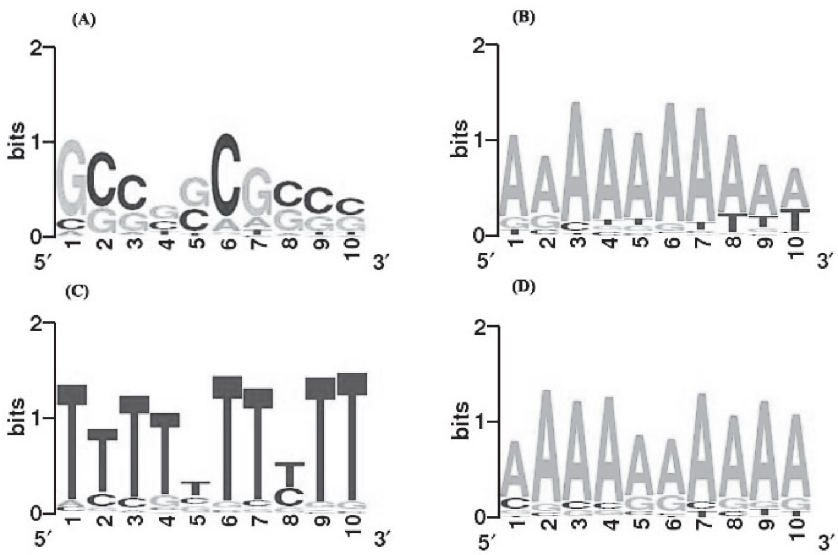

Figure 3 Over-representative *cis*-regulatory elements specific to day-length response. Over-representative *cis*-regulatory element of SD (A) and LD (B) promoters in rice. Over-representative *cis*-regulatory element of SD orthologous (C) and LD orthologous (D) in *Arabidopsis*. The overall height of each stack indicates the sequence conservation at that position (measured in bits), whereas the height of symbols within the stack reflects the relative frequencies of the corresponding nucleic acid at that position.

these genes was determined using a position-analysis program from RSAT. The data revealed that the GC element was predominantly found in SD promoters while the A-rich element was only enriched in LD promoters (data not shown), indicating the specificity of each element in each subset genome.

To further investigate whether the novel *cis*-regulatory motifs specific to day-length in rice (an SD plant) are conserved in *Arabidopsis thaliana* motifs (an LD plant), over-representative motifs in the promoter regions of orthologous genes in *Arabidopsis* were analyzed. The data are shown in Figure 3. Interestingly, the consensus sequences of the over-representative sequences of rice LD promoters and *Arabidopsis* show high homology, suggesting that rice and *Arabidopsis* might have some conserved regulatory mechanisms in response to photoperiod. Thus, to examine the role of the A-rich element, the LD-responsive genes containing this element were identified and are listed in Table 4. Proteins encoded by these predicted genes appear to have diverse biological functions; transcription, metabolism, signal transduction, energy metabolism and protein processes (Table 4), suggesting that the A-rich element plays a role in coordination with various pathways during the LD light response in rice. Moreover, the A-rich element was found in the promoters of floral repressor genes both in rice and *Arabidopsis*, i.e., *AP 2*, *Vip 14*, *FAR 1*, *COP 1* and *ARR1-like* (Mongkolsiriwatana, 2008). FAR 1, COP 1 and ARR1-like proteins are also involved in phyA signaling and circadian rhythm (the major mechanisms of photoperiod control) (Hoecker and Quail, 2001; Lin and Wang, 2004; Hazen *et al*., 2005). In addition, the A-rich element is similar to the CCA1 motif 2 binding site, AAAAAAAATCTATGA (RiceCis-element Searcher, http://hpc.irri.cgiar.org/ tool/nias/ces) which is required by circadian rhythm-controlled genes (McClung, 2006). This evidence supported the hypothesis that the A-rich element is involved in the regulation of phyA and circadian rhythm to inhibit flowering during non-inductive light periods. Thus, this study proposed that the novel A-rich element is one of the key *cis*-regulatory elements in the specific day-length to control flowering in rice.

On the other hand, the over-representative motif (GC element), obtained from the SD promoters of rice is not homologous to the SD promoters of *Arabidopsis*, indicating that this element is unique to rice. It is interesting to further investigate whether the GC element has a specific role in the SD response. The SD-responsive genes containing the GC-elements were identified and are shown in Table 5. Among their known functions, these genes have diverse biological functions, including; stress and defense response, metabolism, protein processing, transcription and signaling. It is possible that the GC element plays a coordinating role in various biological processes during SD light response. The GC element was also found in the promoter of the pathogen resistant genes, *mla1* and *multidrug resistant protein 8*. Although the correlation between plant pathogen and photoperiod has not been reported, pathogen infection is known to accelerate flowering time with more pronounced effects in resistant interaction (Korves and Bergelson, 2003). In addition to pathogen-resistant genes, the promoter of *SLY 1* was found to contain the GC element, suggesting that this element plays a positive role in the regulation of GA signaling. GA is known as a growth regulator to promote flowering. Thus, this result implied that GC element was possibly involved in the regulation of photoperiodic flowering in rice.

It should be noted that most of the SD responsive genes that contain GC elements are unknown genes. These genes might have a specific role in response to SD light in rice. Their functions and regulatory mechanisms might clarify the mystery of different responses to day-length shown by SD and LD plants.

Table 4 List of SD-suppressed genes containing A-rich element in the promoter regions.

TIGR Gene Target	Rice annotation	Functional categories *
LOC_Os01g58310	expressed protein	Unknown
AK106968	expressed protein	Unknown
LOC_Os01g68269	expressed protein	Unknown
LOC_Os08g45200	expressed protein	Unknown
LOC_Os02g10030	hypothetical protein	Unknown
LOC_Os03g02110	expressed protein	Unknown
LOC_Os03g08520	expressed protein	Unknown
LOC_Os05g49630	expressed protein	Unknown
LOC_Os02g50110	expressed protein	Unknown
LOC_Os02g21269	hypothetical protein	Unknown
LOC_Os05g36970	expressed protein	Unknown
LOC_Os05g41630	expressed protein	Unknown
LOC_Os06g11760	expressed protein	Unknown
LOC_Os06g32540	hypothetical protein	Unknown
LOC_Os08g07860	hypothetical protein	Unknown
LOC_Os01g72990	hypothetical protein	Unknown
LOC_Os06g27860	expressed protein	Unknown
LOC_Os04g02880	expressed protein	Unknown
LOC_Os07g02624	expressed protein	Unknown
LOC_Os10g40280	expressed protein	Unknown
LOC_Os08g37660	expressed protein	Unknown
LOC_Os02g52210	RING/C3HC4/PHD zinc finger-like protein	Transcription factor
LOC_Os02g08440	OsWRKY71	Transcription factor
LOC_Os01g64470	Harpin-induced protein 1 containing protein	Transcription factor
LOC_Os08g36920	**ERF (ethylene response factor) subfamily B-1 of ERF/AP2 transcription factor family**	Transcription factor
LOC_Os05g12640	BURP domain	Transcription factor
LOC_Os01g03720	putative myb-related transcription factor	Transcription factor
LOC_Os01g73770	**Transcription factor RCBF2/dehydration and cold-relative**	Transcription factor
LOC_Os10g25230	ZIM motif family protein	Transcription factor
LOC_Os03g28940	pnFL-2, putative, expressed, ZIM motif family protein	Transcription factor
LOC_Os01g74020	**ARR1 protein-like, putative, expressed (LUX)****	Transcription factor
LOC_Os06g44010	WRKY28 protein	Transcription factor
LOC_Os02g36530	helix-loop-helix DNA-binding domain containing protein	Transcription factor
LOC_Os03g07440	putative eribonuclease P, putative, expressed	Transcription factor
LOC_Os03g32220	Zinc finger	Transcription factor
LOC_Os04g23550	**Oryza sativa RERJ1 mRNA for Transcription Factor, PIF 3****	Transcription factor
LOC_Os03g53020	Helix-loop-helix DNA-binding domain containing protein	Transcription factor
LOC_Os01g64310	putative NAC-domain protein	Transcription factor
LOC_Os02g39330	Endochitinase PR4 precursor	Metabolism
LOC_Os10g34840	ripening-related protein 3 precursor, putative, expressed	Metabolism
LOC_Os01g43740	cytochrome P450	Metabolism
LOC_Os08g14190	Sulfotransferase domain containing protein	Metabolism
LOC_Os05g43910	Cytochrome P450 family	Metabolism
LOC_Os06g15410	Aldose 1-epimerase family protein	Metabolism
LOC_04g37430	Lipoxygenase 5	Metabolism
LOC_Os01g72530	EF hand family protein, calmodulin-like protein 41,	Signaling
LOC_Os07g44290	CBL-interacting serine/threonine-protein kinase 1	Signaling
LOC_Os05g07420	putative receptor-like protein kinase (Arabidopsis thaliana)	Signaling
LOC_Os06g49550	**far-red impaired responsive protein** (FAR 1)**	Signaling
LOC_Os01g19050	**far-red impared** (FAR1)**	Signaling
LOC_Os02g54590	Kinase	Signaling
LOC_Os04g42260	Protein phosphatase 2C containing protein, expressed	Signaling
LOC_Os07g35540	putative serine/threonine-specific protein kinase	Signaling
LOC_Os12g19470	Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit	Energy
LOC_Os07g41300	putative NADH dehydrogenase 49kDa protein	Energy
LOC_Os04g51160	AOX1b mRNA for alternative oxidase	Energy
LOC_Os07g02970	**COP 1 interacting protein**	Processing
LOC_Os01g72160	Glutathione S-transferase, putative,expressed	Processing
LOC_Os02g33590	Ubiquitin-protein ligase, putative, expressed	Processing
LOC_Os03g10050	serine O-acetyltransferase, putative	Processing
LOC_Os02g17000	Subtilisin N-terminal Region family protein	Processing
LOC_Os01g51450	nucleosome assembly protein (NAP),	Processing
LOC_Os10g35090	putative membrane-associated protein	Growth/structure
LOC_Os08g23870	Late embryogenesis abundant group 1 family protein	Growth/structure
LOC_Os09g27940	Avr9 elicitor response protein, putative, expressed	Stress/defense
LOC_Os01g03340	Bowman-Birk type bran trypsin inhibitor precursor	Stress/defense
LOC_Os07g05940	**Viviparous-14, neoxanthin cleavage enzyme**	Stress/defense
LOC_Os10g36500	Pectinesterase inhibitor domain	Stress/defense

Note; The flowering time associated genes are highlighted by **bold font**.

* Function categories were classified according to their putative function and consulted with MIPS MATDB database (http://mips.gsf.de/proj/thal/db/index.html).

Table 5 List of SD-induced genes containing a GC element in the promoter regions.

TIGR Gene Target	Rice annotation	Functional categories*
LOC_Os01g34970	multidrug resistance protein 8, putative, expressed	Stress and defense
LOC_Os01g63010	USP family protein, putative, expressed	Stress and defense
LOC_Os12g11990	expressed protein	Unknown
LOC_Os03g54150	expressed protein	Unknown
LOC_Os07g43604	expressed protein	Unknown
LOC_Os06g29730	expressed protein	Unknown
LOC_Os03g61160	expressed protein	Unknown
LOC_Os10g22510	Mla1, putative, expressed	Stress and defense
LOC_Os10g42040	expressed protein	Unknown
LOC_Os04g02640	3-ketoacyl-CoA synthase, putative, expressed	Metabolism
LOC_Os11g20239	expressed protein	Unknown
LOC_Os11g35340	S-adenosylmethionine-dependent methyltransferase	Metabolism
LOC_Os12g04770	Conserved hypothetical protein	Unknown
LOC_Os04g11660	F-box protein interaction domain containing protein	Protein processing
LOC_Os09g13440	expressed protein	Unknown
LOC_Os11g10590	expressed protein	Unknown
LOC_Os01g15910	UTP—glucose-1-phosphateuridylyltransferase family protein	Metabolism
LOC_Os01g39490	hypothetical protein	Unknown
LOC_Os02g54330	F-box domain containing protein , SLEEPY 1	Protein processing
LOC_Os03g47430	hypothetical protein	Unknown
LOC_Os03g47430	hypothetical protein	Unknown
LOC_Os09g08410	hypothetical protein	Unknown
LOC_Os11g34990	hypothetical protein	Unknown
LOC_Os12g30940	F-box domain containing protein	Protein processing
LOC_Os01g60510	hypothetical protein	Unknown
LOC_Os01g41960	retrotransposon protein, putative, unclassified	Unknown
LOC_Os01g72500	retrotransposon protein, putative, unclassified	Unknown
LOC_Os03g07460	retrotransposon protein, putative, unclassified	Unknown
LOC_Os05g40460	retrotransposon protein, putative, unclassified	Unknown
LOC_Os09g36410	retrotransposon protein, putative, unclassified	Unknown
LOC_Os10g28250	retrotransposon protein, putative, LINE subclass	Unknown
LOC_Os02g55250	bHLH transcription factor, putative, expressed	Transcription factors
LOC_Os01g07180	hypothetical protein	Unknown
LOC_Os06g42240	conserved hypothetical protein	Unknown
LOC_Os08g19190	conserved hypothetical protein	Unknown
LOC_Os11g04630	hypothetical protein	Unknown
LOC_Os12g01990	expressed protein	Unknown
LOC_Os08g44670	Calreticulin precursor, putative	signaling
LOC_Os11g12540	hypothetical protein	Unknown

* Function categories were classified according to their putative function and consulted with MIPS MATDB database (http://mips.gsf.de/proj/thal/db/index.html).

CONCLUSION

The common feature of a photoperiod-responsive promoter is the specific combination of LREs, HREs and SREs, showing that the regulatory mechanism in response to day-length is controlled by light in coordination with hormones and stresses. The GARE motif and G-box are the coordinated motifs integrating the GA hormone with photoperiod responses while the MBS element and G-box are the coordinated motifs integrating the ethylene or ABA hormones and stresses with photoperiod. The specific organization of LREs, i.e., G-box, GT-1 motif and GATA motif could be used to identify the phyA-regulated genes from the complexity of photoperiod-responsive pathway. This report showed that phyA is possibly involved in the regulation of the expression of photoperiodic flowering-time genes by the activation of floral inducers or the suppression of floral repressors under SD light via the specific organization of *cis*-regulatory elements in the promoters of target genes. Two novel *cis*-regulatory motifs that are specific to day-length were identified. The novel A-rich element specific to LD light was highly homologous to the element obtained from orthologous genes in *Arabidopsis*. In contrast, the novel GC-element specific to SD light was unique to rice. The A-rich element was involved in the regulation of phyA and circadian rhythm to inhibit flowering while the GC element was involved in the regulation of pathogen resistance and GA signaling to promote flowering. This report has shown for the first time the characteristics of the photoperiod-responsive promoter and elucidated the *cis*-regulatory elements that are specific to the day-length regulation of an SD plant compared to an LD plant. It has provided a new insight into the regulatory mechanism of the photoperiod response at the whole-genome level.

AGKNOWLEDGEMENTS

This work was financially supported by the Corporative Research Network (CRN) in Genetics, Bioinformatics and Bioactive compounds, the Commission on Higher Education. It was partially supported by the Graduate School, Kasetsart University. We wish to thank Associate Professor Dr. Amara Thongpan and Dr. Amorntip Muangprom for useful discussion and critical evaluation of this manuscript.

LITERATURE CITED

Achard, P., H. Cheng, L.D. Grauwe, J. Decat, H. Schoutteten, T. Moritz, D.V.D. Straeten, J. Peng and N.P. Harberd. 2006. Integration of plant responses to environmentally activated phytohormonal signals. **Science** 311: 91-94.

Arias, J.A., R.A. Dixon and C.J. Lamb. 1993. Dissection of the functional architecture of a plant defense gene promoter using a homologous *in vitro* transcription initiation system. **Plant Cell** 5: 485–496.

Babu, M.M., N.M. Luscombe, L. Aravind, M. Gerstein and S.A. Teichmann. 2004. Structure and evolution of transcriptional regulatory network. **Curr. Opin. Struct. Biol.** 14: 283-291

Bagnall, D.J., R.W. King, G.C. Whitelam, M.T. Boylan, D. Wagner and P.H. Quail. 1995. Flowering responses to altered expression of phytochrome in mutants and transgenic lines of *Arabidopsis thaliana* (L) Heynh. **Plant Physiol**. 108: 1495-1503.

Chaboute M.E., B. Clement and G. Phillips. 2002. S phase and meristem-specific expression of tobacco *RNR1 b* gene is mediated by an E2F element located in the 5¢ leader sequence. **J. Biol. Chem**. 277: 17845-17851.

Hazen, S.P., T. F. Schultz, J. L. Pruneda-Paz, J.O. Borevitz, J. R. Ecker and S. A. Kay. 2005.

LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. **Proc. Natl. Acad. Sci. USA** 102: 10387-10392.

Helden, J-v., B. Andre and J. Collado-Vides. 1998. Extracting regulatory site from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. **J. Mol. Biol.** 281: 827-842.

Helden, J-v. 2003. Regulatory sequence analysis tools. **Nucl. Acids Res**. *31*:3593–3596.

Higo, K., Y. Ugawa, M. Iwamoto and T. Korenaga. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database. **Nucl. Acids Res**. **27**:297–300.

Hirano, K., M. Ueguchi-Tanaka and M. Matsuoka. 2008. GID1-mediated gibberellin signaling in plants. **Trends Plant Sci.** 13: 192-199.

Hoecker, U. and P. H. Quail. 2001. The phytochrome A-specific signaling intermediate SPA1 interacts directly with COP1, a constitutive repressor of light signaling in *Arabidopsis*. **J. Biol. Chem.** 276: 38173-38178.

Huang, D., W. Wu, S. R. Abrams and A. J. Cutler. 2008. The relationship of drought-related gene expression in *Arabidopsis thaliana* to hormonal and environmental factors. **J. Exp. Botany. Published online on June 13, 2008.**

Hudson, M. E. and P. H. Quail. 2003. Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data**. Plant Physiol.** 133: 1605–1616.

Izawa, T. 2007. Day length measurement by rice plants in photoperiodic short day flowering. **Int. Rev. Cytol.** 256: 191-222.

Jiao, Y., O.S. Lau and X.W. Deng. 2007. Light-regulated transcriptional networks in higher plants. **Nat. Rev. Genet**. **8:** 217–230.

Johnson, E., M. Bradley, N.P. Harberd and G.C. Whitelam. 1994. Photoresponses of light-grown phya mutants of *Arabidopsis*-phytochrome-A is required for the perception of daylength extensions. **Plant Physiol**. 105: 141-149.

Kęsy, J., B. Maciejewska, M. Sowa, M. Szumilak, K. Kawalowski, M. Borzuchowska and J. Kopcewicz. 2008. Ethylene and IAA interactions in the inhibition of photoperiodic flower induction of *Pharbitis nil*. **Plant Growth Regul.** 55: 43-50.

Korves, T. M. and J. Bergelson. 2003. A developmental response to pathogen infection in *Arabidopsis*. **Plant Physiol.** 133: 339-347.

Kulikowska-Gulewska, H., M. Majewska and J. Kopcewicz. 2000. Gibberellins in the control of photoperiodic flowering transition in *Pharbitis nil*. **Physiol. Plant**. 108: 202-207.

Lenchner, E., P. Achard, A. Vansiri, T. Potuschak and P. Genschik. 2006. F-box protein everywhere. **Curr. Opin. Plant Biol.** 9: 631-638

Lescot, M., P. Déhais, G. Thijs, K. Marchal, Y. Moreau, Y. V. Peer, P. Rouzé and S. Rombauts. 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. **Nucl. Acids Res.** 30: 325-327.

Lin, R. and H. Wang. 2004. *Arabidopsis FHY3/FAR1* gene family and distinct roles of its members in light control of *Arabidopsis* development. **Plant Physiol.** 136: 4010-4022.

McClung, C.R. 2006. Plant circadian clock. **Plant Cell** 18: 792-803.

Menkens, A.E., U. Schindler and A.R. Cashmore**.** 1995. The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. **Trends Biochem. Sci.** 20**:** 506–510.

Mongkolsiriwatana, C. 2008. **Expression Analysis of Photoperiod Responsive Genes in Rice (*Oryza sativa* L.) KDML 105.** PhD thesis. Kasetsart University, Bangkok.

Rabbani, M.A., K. Maruyama, H. Aba, M.A. Khan, K. Katsura, Y. Ito, K. Yoshwara, M.

Seki, K. Shinozaki and K.Yamaguchi-Shinozaki. 2003. Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. **Plant Physiol.** 133: 1755–1767.

Tepperman, J.M., T. Zhu, H-S. Chang, X. Wang and P.H. Quail. 2001. Multiple transcription-factor genes are early targets of phytochrome A signalling. **Proc. Natl. Acad. Sci. USA 98:** 9437–9442.

Venter M. and F.C. Botha. 2004. Promoter analysis and transcription profiling: Integration of genetic data enhances understanding of gene expression. **Physiol. Plant.** 120: 74-83.

Welling, A., P. Kaikuranta and P. Rinne. 1997. Photoperiodic induction of dormancy and freezing tolerance in *Betula pubescens*. Involvement of ABA and dehydrins. **Physiol. Plant.** 100: 119-125

Young, M.W. and S.A. Kay. 2001. Time zones: a comparative genetics of circadian clocks. **Nat. Rev. Genet.** 2: 702-715.

Zhou, J., X. Wang, Y. Jiao, Y. Qin, X. Liu, K. He, C. Chen, L. Ma, J. Wang, L. Xiong, Q. Zhang, L. Fan, and X. W. Deng. 2007. Global genome expression analysis of rice in response to drought and high-salinity stresses in shoot, flag leaf, and panicle. **Plant Mol. Biol.** 63: 591-608.