African Journal of Biotechnology Vol. 10(55), pp. 11477-11482, 21 September, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.1381
ISSN 1684–5315 © 2011 Academic Journals

Full Length Research Paper

Characterization of upstream sequences from the 8S globulin gene of *Vigna radiata*

Yue-Ning Yang, Ying-Fang Niu, Chao Xu, Yin Wang, Jie-Sheng Liu, Wei-Dong Yang and Hong-Ye Li*

Department of Biotechnology, Jinan University, Guangzhou 510632, China.

Accepted 19 August, 2011

Seeds rich in protein in nature, are ideal bioreactors for economic production and storage of value-added recombinant proteins and enzymes for industries. The upstream region of the seed storage protein gene is able to provide an attractive promoter for seed-specific expression of heterologous genes. Our previous research showed that 8S globulin occupied the majority of total soluble protein stored in seeds of mung bean (*Vigna radiata*), a rich source of protein, indicating that the promoter of this gene could be a seed-specific promoter with high activity. To improve the expression of heterologous proteins in plants to act as a bioreactor, the putative seed-specific promoter of 8S globulin gene was characterized in this study. Hence, this potential promoter of beta subunit gene of 8S globulin (*8SGb*) was isolated by genome walking. Analysis with various promoter prediction softwares showed that the promoter sequence possessed many common motifs related to gene transcription in the seed (such as W-box, ABRE element, E-box, etc.). The putative promoter was subsequently cloned into the binary vector pBI121-GFP by replacing the CaMV 35S promoter. The resultant construct designated as pBI-8SGb-GFP was transformed to mung bean cotyledon mesophyll protoplasts. Reporter gene GFP was expressed high in cotyledon protoplasts, which was detected by confocal microscopy, demonstrating the specific activity of *8SGb* promoter in driving gene expression. This study also proved that the 8SGb promoter is an efficient regulatory element for plant seeds to act as a bioreactor.

Key words: Seed-specific, promoter, genome walking, *Vigna radiata*.

INTRODUCTION

Plant bioreactor is popular for its characteristics of low cost, eukaryotic expression and no endogenous pathogen pollution, thus it shares a ground future of marketing as well as commercial value. Though, there are already many useful products such as various types of polypeptide antibody, which were successfully expressed in the plant (De Jaeger et al., 2002), there are also some disadvantages of using plants as bioreactor; like the production yield of the interested heterologous protein is usually low. It has become a main obstacle for commercialization of plants as a bioreactor. To overcome this, developing specific promoters in order to improve the transcription of heterologous recombinant protein can be an effective strategy.

Moreover, considering the stability and accumulation of the expressed heterologous protein in plant cells, a strong promoter seems to be critical for an ideal plant reactor. Nowadays, the most common promoter used in transgenic plants is the cauliflower mosaic virus (CaMV) 35S promoter. It is powerful for expressing heterologous protein possessing several enhancer elements, but as it constitutively expresses and also in every stage of developmental process (Santino et al., 1997) which may deplete nutrient status affecting the plant growth, in addition, some specific product protein which may have negative effect on the plant will occasionally result in death of the plants due to its over accumulation (Romero et al., 1997). Therefore, the aim of our study was to find a novel promoter which is strong, tissue and developmental stage specific in driving the expression of heterologous genes. Recent reports showed that the antibody

*Corresponding author. E-mail: thyli@jnu.edu.cn. Tel: 86-20-85228470.

expressed in grains still maintain activity as high as the original even after three years storage under room temperature. Seeds are rich in proteins in nature, and are ideal bioreactors for economic production and storage of value-added proteins. So, our study focused on identifying seed-specific promoter that can lead the heterologous protein directly to accumulate in seeds which could be a feasible strategy for dealing with the earlier mentioned problem (Li and Jiang, 2006).

Thus, it is quite attractive to concentrate on developing new types of seed-specific strong promoters. Such kind of promoters can drive the target gene to express mainly in seeds, which will help people to regulate the plant metabolic pathways at their will. Further, the nutrient contained in the seeds can also be enriched and enhanced, and the transgenic traits of the plants can be passed on to the next generation under control (Bäumlein et al., 1991; Buchanan et al., 2004). As the use of seed bioreactors in medicine industry has already proved to be efficient nowadays, seed-specific promoters are playing a significant role in the development of seed bioreactors.

In a previous study on mung bean storage proteins, it was found that 8S globulin accumulated in majority of the total soluble protein in seeds, indicating that its gene promoter could be seed-specific with high transcriptional activity. Based on the reported cDNA information (Bemardo et al., 2004), the upstream sequence of beta subunit gene of the 8S globulin was cloned and analyzed by bioinformatics and transient expression. This study potentially provides a high-activity seed-specific promoter that can be used to improve the existing plant bioreactor system.

MATERIALS AND METHODS

Mung bean seed stock was kept in Plant Biotechnology Laboratory in Jinan University. Plants were grown in a growth chamber with 8 h darkness (21℃) and 16 h light (23℃).

DNA extraction and genome walking

Rosette leaves were used for DNA extraction by CTAB genomic DNA extraction method with Universal Genomic DNA Extraction Kit Ver.3.0 (Takara, Dalian, China). Extracted DNA was determined by agarose gel analysis. Then, genome walking was proceeded using three nested specific primers as forward primers which were complement to the 5' upstream sequence of the mung bean 8SG beta subunit cDNA, and four degenerated AP primers in the Genome walking kit (Takara, Dalian, China) in order to amplify the putative promoter of 8SG beta subunit. Forward primers were as follows: Vr72 (5'-AGGAGGGTGTTGGGTTTGGACT-3'), Vr73 (5'-CGGAGTTGAAGTAGAAGGGGTTAT-3') and Vr74 (5'-ACAGAAAGTGATGCCAGGAAAAG-3') according to mung bean 8S globulin beta subunit (*8SGb*) mRNA (GenBank accession EF990626). Amplification procedures were as follows: 1st PCR: 94℃ 1 min; 98℃ 40 s; 94℃ 30 s, 65℃ 1 min, 72℃ 2 min, 5 cycles; 94℃ 30 s, 25℃ 3 min, 72℃ 2 min, 94℃ 30 s, 65℃ 1 min, 72℃ 2 min, 94℃ 30 s, 65℃ 1 min, 72℃ 2 min, 15 cycles; 94℃ 30 s; 44℃ 1 min; 72℃ 2 min; 72℃ 10 min. 2nd and 3rd PCR: 94℃ 30 s, 65℃ 1 min; 72℃ 2 min; 94℃ 30 s; 65℃ 1 min; 72℃ 2 min; 94℃ 30 s; 44℃ 1 min; 72℃ 2 min; 15 cycles; 72℃ 10 min. PCR product was then cloned into pMD19-T vector (Takara) and confirmed by sequence analysis.

Bioinformatic analysis of promoter

The sequencing result was subjected to bioinformatics analysis. Several online tools for promoter prediction as well as BLAST searching were used. PLACE (http://www.dna.affrc.go.jp/database/) was used for transcription elements prediction; while transcription start site and TATA box were predicted with software TSSP (prediction of start of transcription sequences) of Softberry (http://www.softberry.com). Further, NCBI (http ://www.ncbi.nlm.nih.gov) was used for a BLAST search for similar sequences as the upstream sequence of 8SG beta subunit gene.

Expression of GFP reporter in protoplasts of cotyledon

The putative promoter 8SGb was subsequently cloned into the binary vector pBI121-GFP by replacing its original CaMV 35S promoter at the restriction sites of *Hin*dIII-*Bam*HI. The resultant construct was confirmed by sequencing the promoter junctions and designated as pBI-8SGb-GFP. GFP reporter expression was determined by transient expression in protoplasts of cotyledon. Briefly, protoplasts were made from mung bean cotyledon and transformed with the binary vector pBI-8SGb-GFP or pBI121-GFP as the control mediated by PEG-Ca^{2+} method (Yoo et al., 2007). After culture for one night for recovery, protoplasts were observed under a Olympus BX61 upright fluorescence microscope (Olympus,Tokyo, Japan).

RESULTS

Genome walking to obtain upstream sequence of 8SGb

As shown in Figure 1, in the third round of genome walking from the mung bean genomic DNA template, a specific 0.7-kb PCR product was amplified and resolved on the gel, indicating the putative 8SGb promoter sequence. Subsequently, the PCR product was cloned into pMD19-T vector and further confirmed by sequencing analysis.

By analyzing the sequence, it was found to contain 661 bp upstream sequences from the start codon, and it was submitted to GenBank (accession no. GU176353). BLAST search in GenBank revealed two homologous, and both were from the upstream sequences of *Phaseolus vulgaris* genes for beta-phaseolin (GenBank: J01263) and alpha-phaseolin (GenBank: X52626), showing 90% similarity. The regions including the start codon ATG from the three sequences are aligned in Figure 2. Sequences near the start codon showed much higher similarity among the three gene upstream sequences.

The isolated sequence contained many A/T repeats, accounting for more than 60% of the total nucleotides. It was suggested that AT rich fragments are usually the regions where transcription factors bind and thereof regulate the expression of the downstream gene.

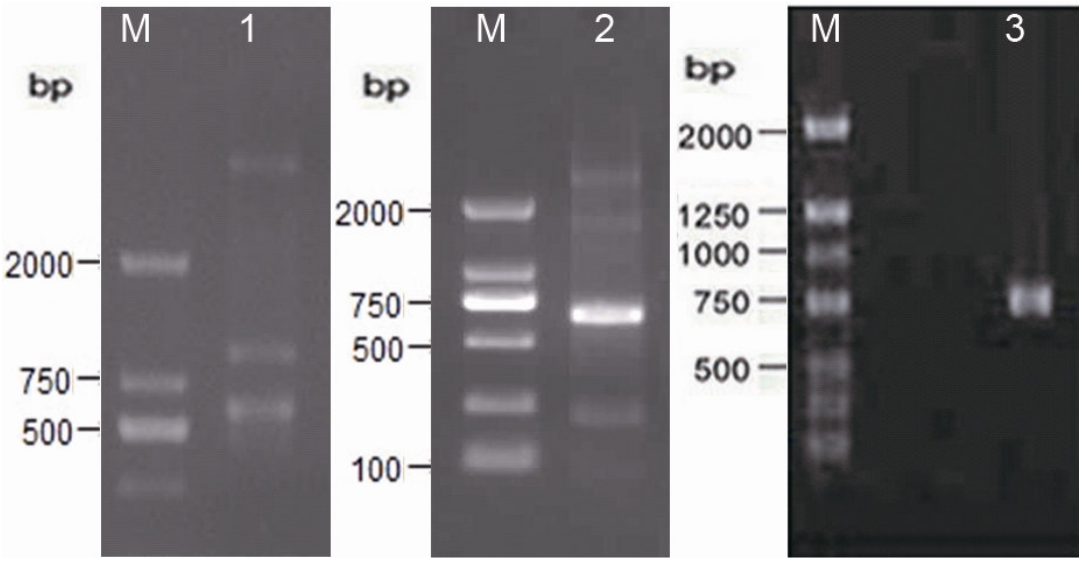

Figure 1. Putative 8SGb promoter isolated by three rounds of genome walking. PCR results were resolved on agarose gel. M, DNA marker; 1, 2 and 3: The PCR results of 1st,2nd and 3rd round of genome walking, respectively.

```
Vr-b ACCACCGGCACGACGATGAACTGCGGCATC--GCCTCGTAGGCAGGCGACTCAGAACCGCCAGCACGCCAGC---ACGCAATCCATCGCAACGGCAACCA 95
Pv-b ----------ATTATTCTACTATGTTTAAATTGTTTTATAGATAGTTTAAAGTAAATATAAGTAATGT----AGTAGAGTGTTAGAGTGTTACCCT 82
Pv-a ----ATTATAATAGGATATTTTGTTCTATTAAAATT---TTATGGAATATATGTAATTCTTAAGAATAATGTTAAGAGTAAATTTTTCAAAATTTATTTG 93

Vr-b AGCCTCAAGCTTGCACAGCACAAACCAC-AAGTGCGTCTCGCACCCGAAACCATTGCCACAGTTGCTTGCCATGACTTT-CAAGGGGTTCCAACGCTTTT 193
Pv-b AAACCATAAAC--TATAACATTTATGGTGGACT--AATTTTCAT-----ATATTTCTTATTGCTTTT-ACCTTTTCTTGGTATGTAAGTCCGTAACTAGA 172
Pv-a ACAAAAAAAATGGTAAAACATATAGTATTAAGTTGATTTTTCGT-----AAAATAGAGACAAATGCA-ACCGCACCCT--TCTTCAATCACACAAATTTG 185

Vr-b CTCACA-CGTGTCACACCGACACT-TCTCCCCTTCATATCAGTCGACTGATATGAAGACCAAAAAAAAAGTATTGTATCCTTTCAAAAACCATGTCCAGA 291
Pv-b ATTACAGTGGGTTGCCATGGCACTCTGTGGTCTTTTGGTTCATGCATGGGTCTTGCG--CAAGAAAAAGACAAAGAACAAAGAAAAAAAGACAAAAC-AGA 269
Pv-a GTT----CAGGTTGTCATGGCACTCTGTAGTCGTTTGGTTCATGCATGGGTCTTACG--CAAGAAAAAGACAAAGAA-------AAAAGCCAAAAC-AGA 271

Vr-b TTCACACAAAATAGAGACGAATGCAACC--G------CGCTGGATGATCAAG-ATGTGGCGTCCACGTATGACAAAA-GCCATGCACACCAACACGTACT 381
Pv-b GAGACAAAACGCAATCACACAACCAACTCAAATTAGTCACTGGCTGATCAAGATCGCCGCGTCCATGTATGTCTAAATGCCATGCAAAGCAACACGTGCT 369
Pv-a GAGA------------------------------------------TCGCCGCGTCCATGTATGTCTAAATGCCATGCACAGCAACACGTGCT 322

Vr-b CAACATGCACTTGAAACACCTAATACAGTTCTTGTTTCTTAGTTCGCACACAGCCAATGCTACTGAGCTACTTCCATCATCACC---AAAACCACATTCT 478
Pv-b TAACATGCACTTTAAATGGCTCACCCA--TCTCAACCCACA-CACAAACACATTGCCTTTTTCTTCATCATCACCACAACCACCTGTATATATTCATTCT 466
Pv-a TAACATGCACTTTAAATGGCTCACC-A--TCTCAACCCACA-CACAAACACAATGCCTTTTTCTTCATCATCACCACAACCACCTGTATATATTCATTCT 418

Vr-b CTTTCGCCACCTCAATTTCTTCATTTCAACACACGTCAACCTCTGTAAATGCGTGTCATCCCATGCCTAAATCTCCATGCATGTTGCAAGGACCT--T-- 574
Pv-b CTTCCGCCACCTCAATTTCTTCACTTCAACACACGTCAACCT--GCATATGCGTGTCATCCCATGCCCAAATCTCCATGCATGTTCCAACCACCTTCTCT 564
Pv-a CTTCCGCCACCTCAATTTCTTCACTTCAACACACGTGAACCT--GCATATGCGTGTCATCCCATGCCCAAATCTCCATGCATGTTCCAACCACCTTCTCT 516

Vr-b -------AATACTTATAAATACCGCTAACATCCCTCATTTCATTCATCATTCATCCA----------------------------ATACCCCAACCCAAC 639
Pv-b CTTATATAATACCTATAAATACCTCTAATATCACTCACTTCTTTCATCATCCATCCATCCAGAGTACTACTACTCTACTACTATAATACCCCAACCCAAC 664
Pv-a CTTATATAATACCTATAAATACCCCTAATATCACTCACTTCTTTCATCATCCATCCATCCAGAGTACTACTACTCTACTACTATAATACCCCAACCCAAC 616

Vr-b TGATATTCAATACTACTATACTATG                                                                          664
Pv-b TCATATTCAATACTACTCTACTATG                                                                          689
Pv-a TCATATTCAATACTACTCTACTATG                                                                          641
```

Figure 2. Alignment of upstream sequences from 8SGb and *Phaseolus vulgaris* genes for beta- and alpha-phaseolin. Vr-b, 8SGb; Pv-b, beta-phaseolin; Pv-a, alpha-phaseolin.

Transcriptional elements predicted by online tools such as PLACE, PlantCare and Softberry revealed that, besides the core element such as TATA box and CAAT box, 8SGb promoter possessed some common elements usually present in seed protein genes, such as Skn-1 motif, RY element, etc (Figure 3), which were

```
ACCACCGGCACGACGATGAACTGCGGCATCGCCTCGTAGGCAGGCGACTCAGAACCGCCAGCACG
                                   ABRE
CCAGCACGCAATCCATCGCAACGGCAACCAAGCCTCAAGCTTGCACAGCACAAACCACAAGTGCG
       CAAT motif                                                E-box
TCTCGCACCCGAAACCATTGCCACAGTTGCTTGCCATGACTTTCAAGGGGTTCCAACGCTTTTCT
       LTR-motif       E-box         W-box
CACACGTGTCACACCGACACTTCTCCCCTTCATATCAGTCGACTGATATGAAGACCAAAAAAAAA
ABRE G-box
GTATTGTATCCTTTCAAAAACCATGTCCAGATTCACACAAAATAGAGACGAATGCAACCGCGCTG
           BoxI
GATGATCAAGATGTGGCGTCCACGTATGACAAAAGCCATGCACACCAACACGTACTCAACATGCA
                        G-box                       G-box     RAV
CTTGAAACACCTAATACAGTTCTTGTTTCTTAGTTCGCACACAGCCAATGCTACTGAGCTACTTC
       E-box                                 CAAT-box
CATCATCACCAAAACCACATTCTCTTTCGCCACCTCAATTTCTTCATTTCAACACACGTCAACCT
       LAMP-element                CAAT-box
CTGTAAATGCGTGTCATCCCATGCCTAAATCTCCATGCAGTGTTGCAAGGACCTTAATACTTATA
        Skn-1 motif          RY-element                    TATA-box
AATACCGCTAACATCCCTCATTTCATTCATCATTCATCCAATACCCCAACCCAACTGATATTCAA
                        +1→
TACTACTATACT
```

Figure 3. The promoter sequence of *8SGb* gene. Transcriptional elements are underlined; transcriptional start site is indicated with +1.

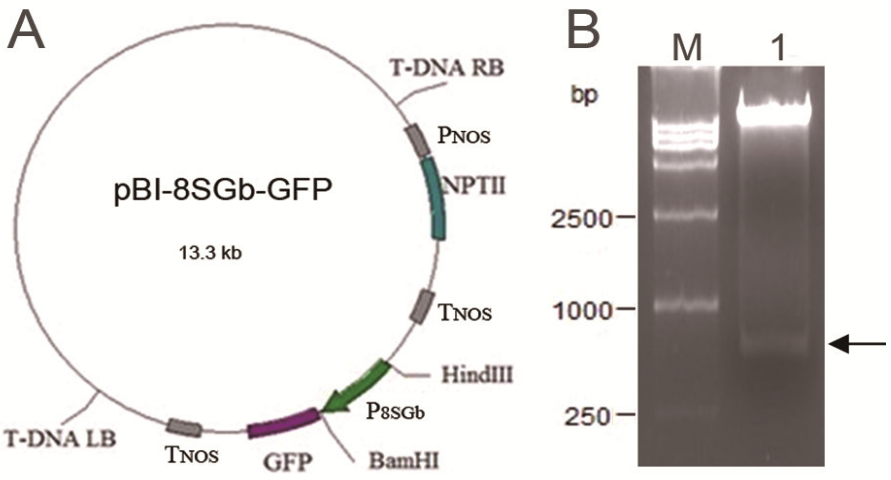

Figure 4. Diagram showing binary vector pBI-8SGb-GFP and restriction digestion analysis. a) P_{8SGb}, Promoter of *8SGb* isolated in this study; P_{NOS}/T_{NOS}, promoter/terminator of *NOS* gene derived from pBI121; b) M, DNA marker; 1, plasmid pBI-8SGb-GFP digested with *Hin*dIII+*Bam*HI, showing a band of released promoter.

demonstrated to have seed-specific promoter activity, and responsible for specific response to environmental factors (Fujiwara and Beachy, 1994).

The 8SGb promoter was then used to substitute CaMV

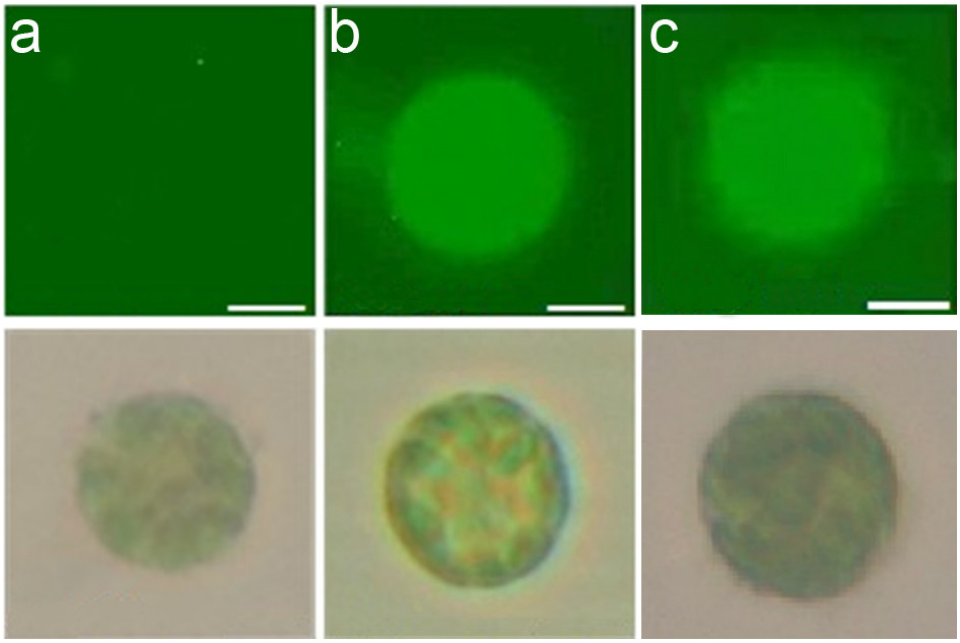

Figure 5. Transient expression of GFP driven by 8SGb promoter in the protoplasts of mung bean cotyledon. Images were captured under fluorescent microscope. Top panel shows the green fluorescence channel and below panel shows the bright view. a) Untransformed control; b) transformed with 8SGb-GFP; c) transformed with 35S-GFP. Bar = 10μm.

35S promoter in plasmid pBI-GFP which was derived from binary vector pBI121 with GUS replaced by GFP. Diagram of the resultant plasmid pBI-8SGb-GFP is shown in Figure 4A, confirmed by restriction analysis (Figure 4B) and sequencing analysis of the promoter.

Subsequently, plasmid pBI-8SGb-GFP was transiently expressed in mung bean cotyledon protoplasts. Since the nature of germinating seeds is still similar to seeds, thus the cotyledon could be used to test seed-specific expression of 8SGb promoter. As shown in Figure 5, GFP reporter driven by 8SGb promoter was observed to be highly expressed in the cotyledon protoplast, which is similar to that under 35S promoter, indicating the relatively high transcriptional activity and seed-specific characteristic of 8SGb promoter.

DISCUSSION

The upstream sequence of *8SGb* gene isolated by genome walking was proved to be a seed-specific promoter containing many promoter specific cis-acting regulatory elements. Expression of GFP reporter further demonstrated the transcriptional activity driven by the promoter sequence. Genome walking was proved to be an efficient approach to amplifying unknown sequences in the genome in this study.

A strong seed-specific promoter is quite beneficial for plant genetic engineering. It is because the strong promoters used today are mainly constitutive promoters, which means that these promoters can continuously drive the expression of heterologous proteins everywhere in the whole plant; the excessive accumulation of heterologous protein may easily lead to abnormality. Moreover, it was revealed that the CaMV 35S promoter, the most common promoter, is a recombination hot spot itself (Koli et al., 1999), which means that gene silencing could be probably triggered. Also, as in some cases, several types of different proteins are often required to be co-expressed in the same organism; various kinds of promoters are needed, considering that too many copies of the same promoter may trigger gene silencing at a higher frequency. Further, the accumulation of heterologous proteins in seeds is an ideal way to store the target proteins because the seed is a stable place for protein storage. It has been reported that antibodies expressed in cereal seeds did not decrease in detectable activity after storing at room temperature for more than three years (Stöger et al., 2000).

As seed specific promoters share lots of advantages when compared with other promoters which may be extensively used nowadays in plant genetic engineering, the cloning and characterization of 8SGb promoter could be a useful alternative and provide an effective tool. So far, our knowledge of this novel seed-specific promoter 8SGb has been limited, which to some extent blocks our capacity to make full use of it. However, from data shown by bioinformatics analysis, there are many environmental responsive elements contained in the sequence. For example, the G-box, E-box and LAMP element are all

sensitive to light and thus might be induced by different intensity of light, while the ABRE and W-box are responsible for factors that can activate the plant defense system such as JA which acts as a signal molecule; or transcriptional factors that express under the induction of pathogen invasion (Mahalingam et al., 2003). The LTR motif can have various responses under different temperature condition. All these lead us to propose that 8SGb promoter could also be inducible by certain environmental factors. Thus, through optimization of the environmental conditions for 8SGb promoter, a higher efficiency of target gene expression could be achieved. At present, the availability of seed-specific promoters is still quite deficient. Development of new seed-specific promoters and in-depth understanding of the promoters between structure and function will be helpful for good regulation of gene expression in genetic engineering. In addition, in the case of multiple-gene transformation, the use of different promoters can accurately regulate gene expression and reduce the gene silencing caused by the use of the same promoter. The use of native plant promoters can also help to avoid transgene silencing which is often associated with the presence of promoters from non-plant origin in the plant genome (Ross et al., 2004; Rubio-Somoza et al., 2006).

In summary, a promoter from seed storage protein gene of mung bean was cloned and characterized in this study, which also provided an efficient regulatory element for plant seed bioreactor.

ACKNOWLEDGEMENTS

This work was supported by Guangdong Natural Science Foundation (No. 20094401120010), and Research Fund for the Doctoral Program of Higher Education of China (No. 9151063101000001).

REFERENCES

Bäumlein H, Boerjan W, Nagy I, Bassüner R, Van Montagu M, Inzé D, Wobus U (1991). A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and Arabidopsis plants. Mol. Gen. Genet. 225: 459-467.

Bemardo AE, Garcia RN, Adachi M, Aangeles JG, Kaga A, Ishimoto M, Utsumi S, Tecson-mendoza EM (2004). 8S globulin of mung bean [*Vigna radiata* (L.) Wilczek]: cloning and characterization of its cDNA isoforms, expression in *Escherichia coli*, purification, and crystallization of the major recombinant 8S isoform. J. Agric. Food. Chem. 52(9): 2552-2560.

Buchanan CD, Klein PE, Mullet JE (2004). Phylogenetic analysis of 5'-noncoding region from the ABA-responsive rab16/17 gene family of soghum, maize and rice provides insight into the composition, organization and function of cis-regulatory modules. Genetics, 168: 1639-1654.

De Jaeger G, Schefer S, Jacobs A, Zambre M, Zobell O, Goossens A, Depicker A, Angenon G (2002). Boosting heterologous protein production in transgenic dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. Nat. Biotechnol. 20(12): 1265-1268.

Fujiwara T, Beachy RN (1994). Tissue-specific and temporal regulation of a beta-conglycinin gene: roles of the RY repeat and other cis-acting elements. Plant Mol. Biol. 24(2): 261-272.

Kohli A, Griffiths S, Palacios N, Twyman RM, Vain P, Laurie DA, Christou P (1999). Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hot spot in CaMV 35S promoter and confirms the of microhomology mediated recombination. Plant J. 17: 591-601.

Li HY, Jiang L (2006). Seeds as bioreactors for producing biopharmaceuticals. In Floriculture, Ornamental and Plant Biotechnology: advances and topical issues (1st ed). Jaime A. Teixeira da Silva, ed. Global Sci. Books. pp. 39-45.

Mahalingam R, Gomez-Buitrago A, Eckardt N, Shah N, Guevara-Garcia A, Day P, Raina R, Fedoroff NV (2003). Characterizing the stress/defense transcriptome of Arabidopsis. Genome Biol. 4(3): p. 20.

Romero C, Bellés JM, Vayá JL, Serrano R, Culiáñez-Macià FA (1997). Expression of the yeast trehalose-6-phosphate synthase gene in transgenic tobacco plants: pleiotropic phenotypes include drought tolerance. Planta, 201(3): 293-297.

Ross EJ, Stone JM, Elowsky CG, Arredondo-Peter R, Klucas RV, Sarath G (2004). Activation of the *Oryza sativa* non-symbiotic haemoglobin-2 promoter by the cytokinin-regulated transcription factor, ARR1. J. Exp. Bot. 55(403): 1721-1731.

Rubio-Somoza I, Martinez M, Diaz I, Carbonero P (2006). HvMCB1, a R1MYB transcription factor from barley with antagonistic regulatory functions during seed development and germination. Plant J. 45(1): 17-30.

Yoo SD, Cho YH, Sheen J (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat. Protocol. 2: 1565-1572.

Santino CG, Stanford GL, Conner TW (1997). Developmental and transgenic analysis of two tomato fruit enhanced genes. Plant Mol. Biol. 33(3): 405-416.

Stöger E, Vaquero C, Torres E, Sack M, Nicholson L, Drossard J, Williams S, Keen D, Perrin Y, Christou P (2000). Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. Plant Mol. Biol. 42: 583-590.