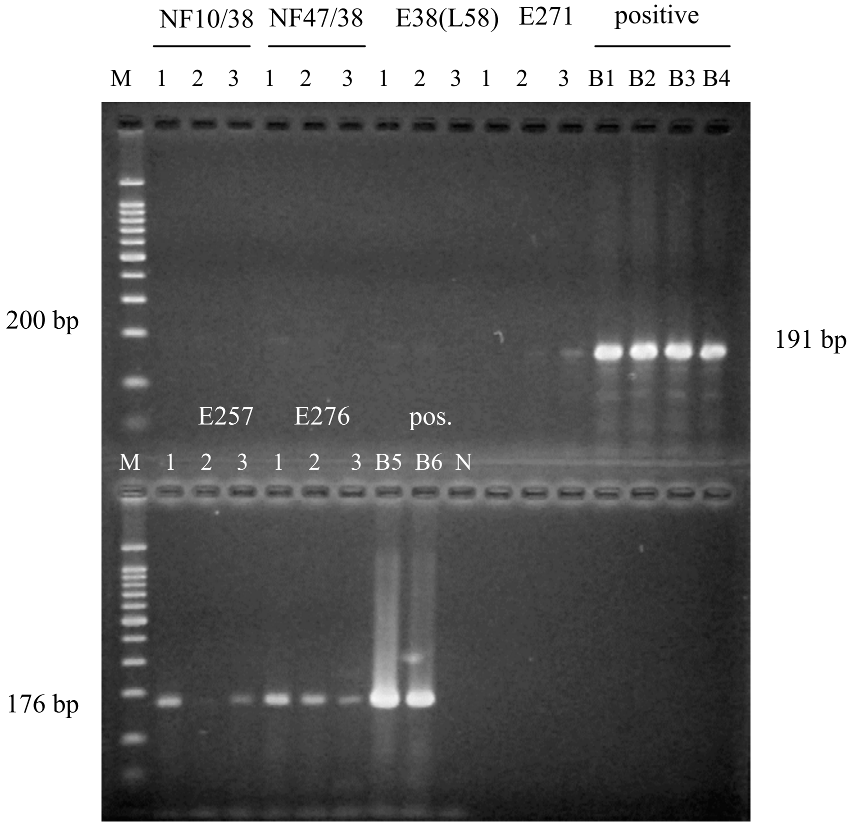

Fig.7.1. 3% agarose gel analysis of PCR-amplified products of the variable regions of flagellin genes of *B. pseudomallei* and *B. thailandensis* from soil suspension. *The upper gel*: PCR-amplified products of *B. pseudomallei*, NF10/38, NF47/38, E38(L58), E271 inoculated in soils 1, 2, and 3 in lanes 1, 2, and 3 respectively. Lane M is a 100 bp ladder DNA marker. Lanes B1, B2, B3, and B4 are PCR of cultured broth of NF10/38, NF47/38, E38(L58), and E271 respectively (positive controls). *The lower gel*: PCR-amplified products of *B. thailandensis*, E257 and E276 inoculated in soils 1, 2, and 3 in lanes 1, 2, and 3 respectively. Lane M is a 100 bp ladder

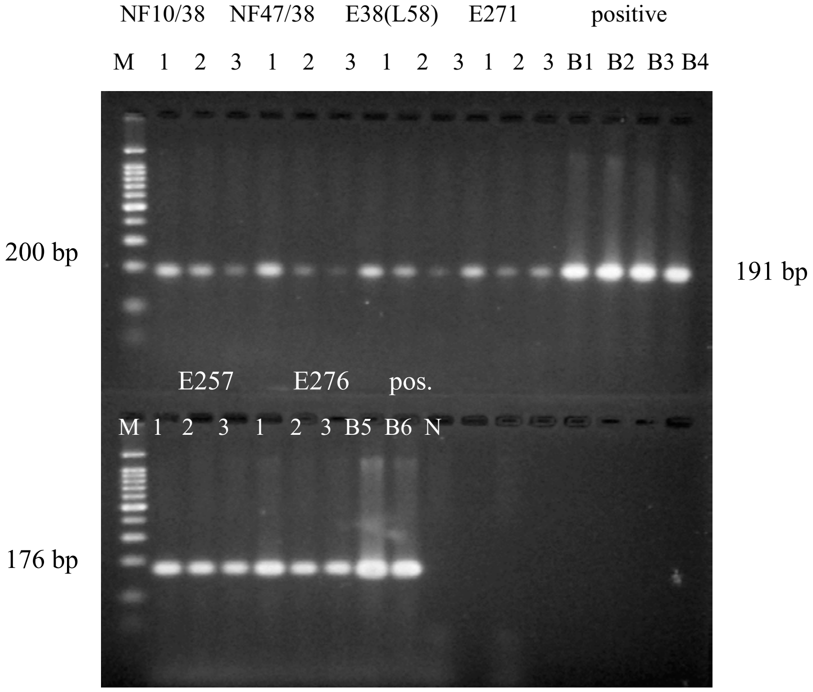

Fig.7.2. 3% agarose gel analysis of PCR-amplification of variable regions of flagellin genes of *B. pseudomallei* and *B. thailandesis* from soil suspensions by addition of BSA in PCR reaction. *The upper gel*: PCR-amplified products of *B. pseudomallei*, NF10/38, NF47/38, E38(L58), E271 inoculated in soils 1, 2, and 3 in lanes 1, 2, and 3 respectively. Lane M is a 100 bp ladder DNA marker. Lanes B1, B2, B3, and B4 are PCR of cultured broth of NF10/38, NF47/38, E38(L58), and E271 respectively (positive controls). *The lower gel*: PCR-amplified products of *B. thailandensis*, E257 and E276 inoculated in soils 1, 2, and 3 in lanes 1, 2, and 3 respectively. Lane M is a 100 bp ladder DNA marker. Lanes B5, and B6 are PCR of cultured broth of 257 and E276 (positive controls). Lane N is a negative control of PCR run.

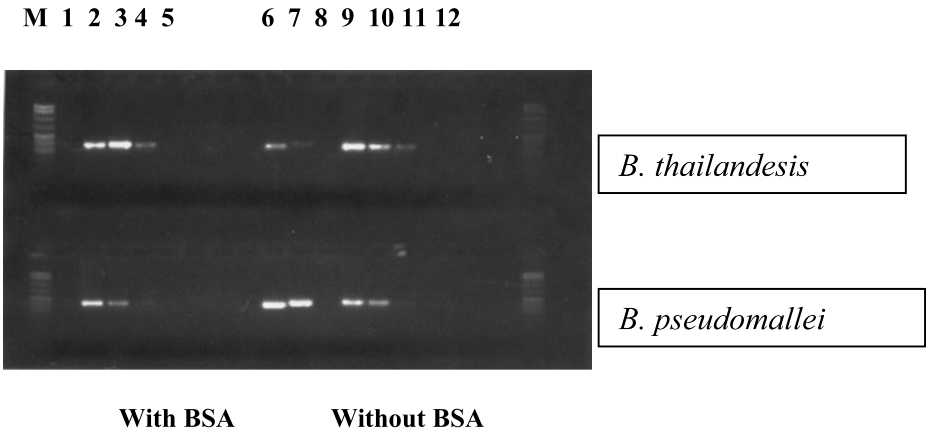

Fig. 8. 3% agarose gel analysis of PCR-amplification of variable regions of flagellin genes of *B. thailandesis* (upper) and *B. pseudomallei* (lower) from soil suspensions by with BSA in lanes 1, 2, 3, 4, 5 using 0.5 M primers and 10^{7}, 10^{6}, 10^{5}, 10^{4}, 10^{3}, respectively. ; and without BSA in lanes 8, 9, 10, 11, 12. using 0.5 M primers and 10^{7}, 10^{6}, 10^{5}, 10^{4}, 10^{3} respectively. Lane M is a 100 bp ladder DNA marker. Lanes 6, and 7 are PCR of cultured broth of Ara+ and Ara- (positive controls). Lane N is a negative control of PCR run.

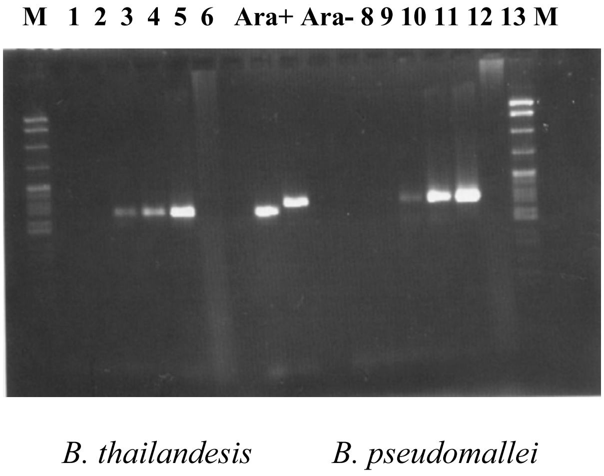

Fig. 9. Sensitivity detection of *B. thailandesis* and *B. pseudomallei* from soil suspensions by without BSA in lanes 1 and 8, 2 and 9, 3 and 10, 4 and 11, 5 and 12 and 6 and 13 using 0.5 M primers and bacterial cells of 10^{1}, 10^{2}, 10^{3}, 10^{4}, 10^{5} and 10^{6} respectively. Lane M is a 100 bp ladder DNA marker. Lanes Ara+, and Ara- are PCR of cultured broth of Ara+ and Ara- (positive controls).

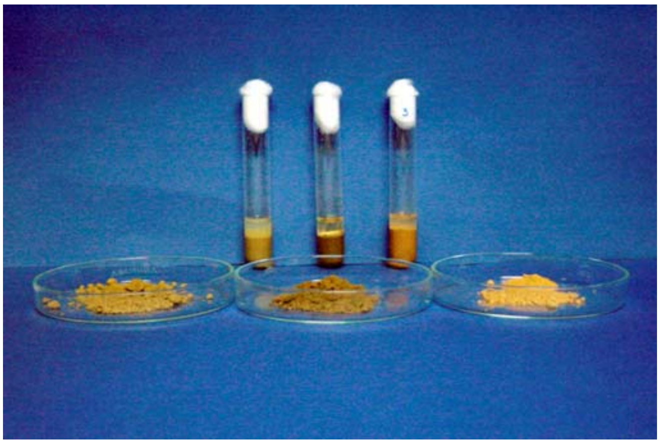

Fig.10. Color of soils; soils collected from rice field in Hoi Kha Yung, Ubon Ratchatani (soil 1), rice field in Varinchamrab, Ubon Ratchatani (soil 2) and Song Khla University, Song Khla (soil 3).

CLONING AND CHARACTERIZATION OF A LIGHT-INTENSITY-RESPONSIVE GENE OF *SYNECHOCOCCUS* PCC7942.

***Wipa Chungjatupornchai*, Prae Plansangkate, Siam Popluechai, Sirirat Fa-aroonsawat, Petcharin Pareong and Sakol Panyim**

Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakornpathom 73170, Thailand.

Abstract

Promoter-active fragments of *Synechococcus* PCC7942 were isolated by transcriptional gene fusion to the promoterless *β-glucuronidase* (*GUS*) gene of *E. coli*, which was used as a reporter gene. One of the isolated strong promoter-active fragments, designated E10, was regulated by light intensity and not functioned in *E. coli*. Nucleotide sequence at 3'end of the E10 fragment harbored the 3' end of the *htpG* heat shock protein gene. Deletion analysis revealed that the light-intensity-responsive promoter was located at the 3'end and downstream of the *htpG* gene. In order to identify the gene regulated by the light-intensity-responsive promoter, the downstream sequence of *htpG* gene was isolated using genomic walking method. Nucleotide sequence analysis revealed two possible ORF at downstream of the *htpG* gene: one encoding 76 amino acids (ORF76) and located at the same orientation as *htpG* gene, the other encoding 315 amino acids of the C-terminal of a protein and located at the opposite orientation. No significant similarity to the ORF76 and the 315 deduced amino acid sequence was detected in the database. The *ORF76* transcript was detected using RT-PCR method. Thus, the *ORF76* was expressed in *Synechococcus.* The RT-PCR of *ORF76* under high light intensity was two to three folds higher than that under low light intensity. These results agreed well with that of E10-*GUS* transcript. Thus, the *ORF76* was under the control of E10 promoter-active fragment. Differentiation of the 5' end of *ORF76* transcripts of wild type *Synechococcus* revealed the true and processed sites designated +1T and +1P, respectively. The +1P site is identical to the transcription initiation site of the *E10-GUS* transcripts determined by the 5'RACE-PCR method. The stem-loop structure at nt.-9 to +63, which included the +1T and +1P sites of *ORF76* transcripts, was probably an signal for processing of the 5'UTR of *ORF76*. Downstream sequence of this stem-loop did not affect the processing of the 5'UTR of *ORF76*, since the transcription initiation site of *E10-GUS* detected using 5'RACE-PCR method was identical to the +1P site of *ORF76*. *Synechococcus* strain totally lacking chromosome copies of *ORF76* was constructed. The *ORF76* was not essential to cell viability but lacking *ORF76* might affect cell growth at high temperature. The phenotypes of the *ORF76* mutant strain are under investigation.

Key words: cyanobacteria, *Synechococcus* PCC7942, β-glucuronidase, light-intensity-responsive gene.

Introduction

Cyanobacteria (blue green algae) are photosynthetic and easily proliferate in tropical aquatic habitats. Cyanobacteria have become hosts of interest for expression of the high value proteins, since its ability to grow in a minimal media with light offers a possibility of low cost production. Cyanobacteria have been used as hosts to express several heterologous genes. For example, attempts have been made to express the mosquitocidal protein genes of *Bacillus sphaericus* and *Bacillus thuringiensis* subsp. *israelensis* in order to provide an alternative biological insecticide for control of mosquito populations [1, 2, 3]. They have also been used in the expression of salmon growth hormone gene in order to produce a feed additive for fish [4], or the expression of ethylene-forming enzyme in order to exploit atmospheric CO_2 as a substrate [5]. However, the level of heterologous gene expression in cyanobacteria is low when compared with that in *E. coli*. A possible way to improve the gene expression is to use an endogenous strong promoter. However, current knowledge of the structure and function of promoters recognized within cyanobacteria is still limited. Although cyanobacteria are classified as eubacteria [6], their RNA polymerase holoenzyme is unique in that it contains a subunit γ, in addition to the $\alpha_2\beta\beta'\sigma$ structure common to the RNA polymerase of other eubacteria [7]. The genetics of cyanobacteria is highly heterogeneous, the GC content of their DNA ranges from 35% to 71% which is as varied as the whole kingdom of bacteria [8]. Our study showed that some strong promoters in *E.coli* may become weak promoters in cyanobacteria [9]. Cyanobacterial promoters may be different from *E.coli* promoters [10].

In order to isolate cyanobacterial promoters for structure and functional studies, the unicellular *Synechococcus* PCC7942 (previously referred to as *Anacystis nidulans* R2, Pasteur Culture Collection no. 7942) was used, because it is both physiologically and genetically well characterized and highly transformable [11]. We isolated several *Synechococcus* PCC7942 strong promoters [10]. Promoter-active fragments of *Synechococcus* PCC7942 were isolated by transcriptional gene fusion to the promoterless *β-glucuronidase* (*GUS*) gene of *E. coli*, which was used as a reporter gene. Several of the isolated promoter-active fragments expressed GUS activity in *Synechococcus* comparable with that of the λP_R promoter. Only 10% of the isolated promoter-active fragments also functioned in *E. coli*.

One of the isolated strong promoter-active fragments, designated E10, was regulated by light intensity [10]. However, the gene regulated by the light-intensity-responsive promoter, E10, was unknown. Thus, cloning and characterization of DNA downstream of the E10 promoter will elucidate the gene that is regulated by light intensity. In this study, the downstream sequence of the E10 fragment was isolated and characterized. The true and processed transcription initiation sites of the *ORF76* transcripts, possibly regulated by the E10 promoter, were differentiated. The *ORF76* inactivated strain was constructed. The phynotypes of the ORF76 inactivated strain is under investigation.

Materials and Methods

Isolation of DNA sequence downstream of the E10 promoter-active fragment.

Isolation of DNA sequence downstream of the E10 was carried out following the manufacture's manual (Universal Genomic Walker kit, CloneTech, USA). In brief, chromosomal DNA of *Synechococcus* PCC7942 R2Spc isolated as previously described [12] was digested with PvuII restriction enzyme. The purified PvuII digested DNA was ligated to Genomic Walker Adaptor (provided with the kit). The adaptor ligated DNA was used as template for PCR carried out using Advantage Genomic PCR kit (CloneTech, USA) with the Adapter Primer 1 (AP1, provided with the kit) and primer Dwn-htpG1:

5'-AGGGGCTTGGGCTGAGAGCGATTGC-3' (located at 3'end of the *htpG* gene, based on sequence accession number AB010001) at 94°C for 1 min, followed by 35 cycles of 94°C for 30 sec, 68°C for 3 min and terminate after the last cycle at 68°C for an additional 12 min. The PCR products were electrophoresed in 1% agarose gel. The PCR product of expected size was excised from the gel, purified and cloned into pGEMT-easy vector (Promega).

Nucleotide sequence analysis

DNA sequences of PCR products cloned in pGEMT-easy were determined by automated sequence analyzer (Perkin Elmer, ABI, Model 377). Sequence comparison and analysis were carried out with the Wisconsin Package Version 9.1 software, Genetic Computer Group (GCG, Madison, WI, USA) and the BLAST program [13] with the GenBank database and the genome database for *Synechocystis* PCC6803 [14].

Detection of ORF76 and GUS mRNAs using relative reverse transcription (RT)-PCR

The cDNA of *ORF76* or *GUS* transcripts were co-synthesized with cDNA of *16S* rRNA transcripts from 5 μg of DNAse I- treated RNA with primers ORF76-rev or GUS1, 16S-rev2 and SuperScript II RNAse H-free Reverse Transcriptase (Life Technologies). Two microliters of the reverse transcription reaction mixture were then used as a template for PCR. The PCR products of ORF76 and *GUS* cDNA were obtained by co-amplification with *16S* cDNA, at 94°C for 1 min, 57°C for 30 s and 72°C for 1 min. The PCR of *ORF76* and *GUS* were amplified for 25 cycles, whereas that of *16S* was for 10 cycles. Primers ORF76-fwd and ORF76-rev, or GUS-fwd and GUS1 were added at cycle 1, whereas primers 16S-fwd and 16S-rev2 were added at cycle 16. After PCR amplification, the reaction mixture was resolved in a 2% agarose gel containing ethidium bromide. An image of the gel was captured with UVP (Life Sciences, UK).

Analysis of ORF76 transcription initiation sites

The procedure used to differentiate genuine initiated transcripts from processed transcripts was carried out essentially as described [15]. In brief (see Figure. 6A), 10 μg total RNA from *Synechococcus* wild type was treated with 30 units of tobacco acid pyrophosphatase (TAP) (Epicenter Technologies, USA) or incubated without TAP in the presence of 80 units of RNasin (Promega) for 90 min at 37°C, extracted with phenol/chloroform, then precipitated with 0.1 volume of 3 M

NaOAc, pH 5.2 and 2.5 volumes of EtOH. Pellets were rinsed with 70% EtOH in DEPC treated dH_2O and then resuspended in 58 µl of DEPC.dH_2O and combined with 10 µl of 10X buffer, 4 µl (160 units) of RNasin, 10 µl of dimethyl sulfoxide, 12 µl (240 units) of RNA ligase, and 6 µl (150 ng) of the 35-mer RNA oligonucleotide. Samples were incubated for 60 min at 37°C and then extracted with phenol/chloroform. The treated RNA was annealed with primer ORF76-rev1 and the cDNA was synthesized with SuperScript II RNAse H-free Reverse Transcriptase (Life Technologies). First PCR was carried out with primers ORF76-rev6 and ESK for 35 cycles at 94°C for 1 min, 55°C for 30 s and 72°C for 1 min. One microliter of a 20-fold dilution of the first PCR product was then subjected to nested PCR. The conditions of nested PCR were identical to those for first PCR except that the primer ORF76-rev5 were used in place of primer ORF76-rev6. The nested PCR products were then cloned into pGEMT-easy vector (Promega). DNA sequences of PCR products were determined by automated sequence analysis (Perkin Elmer, ABI, Model 377).

The mRNA secondary structure prediction

The secondary structures of upstream and downstream regions of *ORF76* were analyzed using the computer program mfold version 3.1 [16, 17].

Inactivation of ORF76 in Synechococcus PCC7942

The procedure used to inactivate ORF76 in *Synechococcus* was carried out essentially as described in Figure 4. The PCR product of ORF76 amplified with primers E10fok and ORF-rev3 was cloned into plasmid pGEMT-easy, resulted in pGEMT-ORF76. In order to inactivate the ORF76 of pGEMT-ORF76, the 2 kb SmaI fragment harboring the Sm^R/Spc^R gene of pHP45Ω [18] was cloned into the fill-in AvaI site of ORF76. The resulting plasmid pGEMT-ORF76Ω was linearized with ScaI and transformed into *Synechococcus*. By homologous recombination, the ORF76 on some copies of *Synechococcus* chromosomes was replaced by the Sm^R/Spc^R gene. The transformants were segregated to obtain the homologous mutant strain.

Results and Disussion

The E10 promoter-active fragment

It has been shown that the E10 promoter-active fragment was regulated by light intensity and not functioned in *E. coli*. GUS activities of E10 in *Synechococcus* PCC7942 grown at 1,500 and 5,000 Lux were 65 (±7) and 298 (±30) nmol/min/mg protein, respectively [10]. Nucleotide sequence at 3'end of the E10 fragment immediately upstream of the promoterless-*GUS* gene showed that it harbored the 3' end of the *htpG* heat shock protein gene of *Synechococcus* PCC7942 (see Figure 1). Deletion analysis revealed that the light-intensity- responsive promoter was located at the 3'end and downstream of the *htpG* gene[Chungjatupornchai, et al, unpublished data]. Thus it was expected that the gene regulated by the light-intensity-responsive promoter may locate at downstream of the *htpG* gene.

The two open reading frames downstream of the htpG gene

The downstream sequence of E10 fragment was isolated using the genomic walking method. PCR product of genomic walking library harboring the downstream sequence of E10 was cloned into the pGEMT-easy for further nucleotide sequence identification (Figure 2).

Nucleotide sequence analysis revealed that there were two possible open reading frames (ORF) at downstream of the *htpG* gene (Figure 3). One ORF encoding 76 amino acids (designated ORF76) was located at the same orientation as *htpG* gene, whereas, the other encoding 315 amino acids of the C-terminal of a protein was located at the opposite orientation. Therefore, it is possible that the ORF76 is regulated by the light-intensity- responsive promoter. No significant similarity to the ORF76 and the 315 deduced amino acid sequence was detected in the database. Thus, they are probably hypothetical proteins. However, sequence analysis of the 315 deduced amino acids revealed that it contained the PAS, GGDEF and DUF1 conserved domains (Figure 3 and 4). The PAS domains (accession no. smart00091) appearing in eubacteria and eukaryotes are important signaling modules that monitor changes in light, redox potential, oxygen, small ligands and overall energy level of a cell. The GGDEF domain (accesssion no. PF00990) is found linked to a wide range of non homologous domains that containing bacterial signaling domains. The function of GGDEF domain is unknown. However, it may be a phosphoesterase based on the pattern of conserved Asp and His residues. The DUF1 domain (accession no. smart00091) is apparently occurred exclusively in bacteria. Its function is unknown. It is likely to participate in prokaryotic signaling process.

The expression of ORF76 in Synechococcus PCC7942

In order to investigate whether the *ORF76* is expressed and regulated by the light-intensity responsive promoter, relative RT-PCR method was used to detect the *ORF76* transcript (Figure 5). Result showed that transcript of ORF76 was detected. Thus, the *ORF76* is expressed in *Synechococcus*. The RT-PCR of *ORF76* under high light intensity was two to three folds higher than that under low light intensity. These results agree well with that of E10-*GUS* transcript. Thus, the ORF76 may be under the control of E10 promoter-active fragment.

Differentiation of true from processed transcription initiation sites of ORF76 trnascripts

The transcription initiation site of the *E10-GUS* transcripts determined by the 5’RACE-PCR method was shown in Figure. 1. Chromosomal DNA sequence analysis revealed that downstream of E10 promoter-active fragment contained the ORF76 (Figure 3). To differentiate the true from processed transcription initiation sites of *ORF76* transcripts, we performed the experiments as described in Figure 6A. Treatment of total RNA with TAP converts all 5’ triphosphates to monophosphate and comparison of the effect of TAP treatment to untreated samples should reveal transcripts that had 5’ triphosphates. Figure 6B shows the PCR products generated by amplification of *GUS* cDNAs synthesized from TAP treatment and untreated total RNA of wild type *Synechococcus*. A distinct band of 200 bp was obtained in TAP-treated samples that was not detected in untreated samples (Figure 6B), indicating the presence of true transcripts. The band of 170 bp that appeared in both TAP-treated and untreated samples (Figure 6B), indicates the presence of processed transcripts. The two PCR products from TAP-treated sample were cloned and the DNA sequences were determined. Sequencing analysis of independent clones harboring the 200-bp PCR product revealed that the true transcription initiation site (designated +1T) was located in E10 fragment upstream of *ORF76* (Figure 1 and 7), whereas the independent clones harboring the 170-bp PCR product revealed that the processed transcription initiation site (designated +1P) was located downstream of the +1T site. The +1P site is identical to the transcription initiation site of the *E10-GUS* transcripts determined by the 5'RACE-PCR method.

Processing of 5' untranslated region (UTR) of ORF76 transcripts is probably signalled by stem-loop structure

The predicted secondary structure of upstream region of *ORF76* revealed three stem-loop structures at nt -81 to –60, nt.-48 to –18 and nt.-9 to +63 with the predicted ΔG value of -7.3, -10.3 and –28.7 kcal/ mol, respectively (Figure 7). The stem-loop at nt.-48 to –18, which included the stop codon of *htpG* gene, was probably the transcription terminator of the *htpG* gene. The stem-loop at nt.-9 to +63, which included the true and processed transcription initiation sites of *ORF76* transcripts, was probably an signal for processing of the 5'UTR of *ORF76*. Downstream sequence of this stem-loop (nt.-9 to +63) did not affect the processing of the 5'UTR of *ORF76*, since the transcription initiation site of *E10-GUS* detected using 5'RACE-PCR method was identical to the processed site of *ORF76.* The stem-loop at nt.+295 and +332 located immediately after the stop codon of *ORF76* gene was probably the transcription terminator (Figure 7).

The essence of ORF76 in Synechococcus PCC7942

To construct *Synechococcus* stain totally lacking chromosome copies of *ORF76*, the Sm^R/Spc^R gene was used to inactivate *ORF76* (Figure 8). The resulting mutant strains were confirmed using PCR. Figure 9 shows the PCR products generated by amplification of chromosomal DNA with primers E10fok and ORF-rev3. The PCR products with the expected size of 550 bp and 2.55 kb were detected in wild type and mutant strains, respectively. The result indicates that mutant strains totally lack chromosome copy of *ORF76.* Southern blot analysis of the PCR products and chromosomal DNA also conformed that mutant strains totally lack chromosome copy of *ORF76* (see Figure 9). Thus, *ORF76* is not essential to cell viability. Preliminary results indicated that lacking *ORF76* might effect cell growth at high temperature. The phenotypes of the *ORF76* mutant strain are under investigation.

Acknowledgments

This work was supported in part by The Thailand Research Fund (TRF). W. Chungjatupornchai was a recipient of TRF. P. Plansangkate is supported by The Ph.D-MD Program of Mahidol University. S. Popluechai is supported by Development and Promotion of Science and Technology Talents Project of Thailand (DPST). S. Panyim is a TRF Senior Research Scholar.

Reference:

1. Angsuthanasombat C, Panyim S (1989) Biosynthesis of 130-kilodalton mosquito larvicide in the cyanobacterium *Agmenellum quadruplicatum* PR-6. Appl Environ Microbiol **55**: 2428-2430.
2. Chungjatupornchai W (1990) Expression of the mosquitocidal-protein genes of *Bacillus thuringiensis* subsp. *israelensis* and the herbicide-resistance gene *bar* in *Synechocystis* PCC6803. Curr Microbiol **21**:283-2883.
3. Tandeau de Marsac N, de la Torre F, Szulmjster J (1987) Expression of the larvicidal gene of *Bacillus sphaericus* 1593M in the cyanobacterium *Anacystis nidulans* R2. Mol Gen Genet **209**: 396-398.
4. Kawata Y, Yamano N, Kojima H, Itoh S (1991) Expression of salmon growth hormone in the cyanobacterium *Agmenellum quadruplicatum*. Biotech Lett **13**: 851-856.
5. Fukuda H, Sakai M, Nagahama K, Fujii K, Matsuoka M, Inoue Y, Ogawa T (1994) Heterologous expression of the gene for the ethylene-forming enzyme from *Psuedomonas syringae* in the cyanobacterium *Synechococcus*. Biotech Lett **16**: 1-6.
6. Giovannoni SJ, Delong EF, Olsen GJ, Pace NR (1988) Evolutionary relationships among cyanobacteria and green chloroplasts. J Bacteriol **170**:720-72.
7. Schneider GJ, Haselkorn R (1988) RNA polymerase subunit homology among cyanobacteria, other eubacteria, and archaebacteria. J Bacteriol **170**: 4136-4140.
8. Rippka R (1988) Recognition and identification of cyanobacteria. Methods Enzymol **167**: 28-67.
9. Chungjatupornchai W. and Panyim S. (1994) Advances in heterologous gene expression in cyanobacteria. In Phang, et. al. (eds), Algal Biotechnology in the Asia-Pacific Region. University of Malaya, Kuala Lumpur, Malaysia. p.186-189.
10. Chungjatupornchai W., Senawong T. and Panyim S. (1999) Isolation and characterization of *Synechococcus* PCC7942 promoters; tRNA^{PRO} gene functions as a promoter. Curr. Microbiol. **38**: 210-216.
11. Porter RD (1986) Transformation in cyanobacteria. Crit Rev Microbiol **13**: 111-132
12. Draper J, Scott R (1988) The isolation of plants nucleic acids. In: Draper J,Scott R. Armitage P. Walden R (eds) Plant genetic transformation and gene expression: a laboratory manual. London: Blackwell Scientific Publications, pp 199-236.
13. Altschul SF, Gish W, Myer EW, Lipman DJ (1990) Basic local aligment search tool. J Mol Biol **215**: 403-410.
14. Nakamura Y, Kaneko T, Hirosawa M, Miyajima N, Tabata S (1998) CyanoBase, a www database containing the complete nucleotide sequence of the genome of *Synechocystis* sp. Strain PCC6803. Nucleic Acids Res **26**: 63-67.
15. Chungjatupornchai W, Fa-aroonsawat S and Panyim S (2002) Characterization of regions of the cyanobacterial tRNA^{pro} gene that affect the expression of a β-glucuronidase reporter gene. FEMS Microbiology Letters **211**: 57-64.
16. Zuker, M., Mathews, D.H. and Turner, D.H. (1999) Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide in RNA Biochemistry and Biotechnology, (Barciszewski, J. and Clark, B.F.C., eds.), pp. 11-43. NATO ASI Series, Kluwer Academic Publishers.
17. Mathews, D.H., Sabina, J., Zuker, M. and Turner, D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. **288**: 911-940.
18. Prentki, P and Krisch, HM (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. Gene **29**: 303-313.

E10 //

1 CCCAGTCAGT AGTCAGAGAG CAATACCGTT AAACCCTGCC CACGCGGGGT

51 TTTTTTATGG GCGATCGCAG ACGCGTCCGT AAGATGAAGG GGTGATGTTG

Sau3A1

101 GCCATCGCCT CAATGCT**GAT CCGGCTAATA TACCCGATGA CCATGTCCTG**

Sau3A1

151 **CTCGTGAATA CCGCGCATCC TCTTGTCCAG AACATCCTTA GTCTGCAGCA**

201 **GGGGGCTATC CTCAGCAGTG ATGGACACTC TCCAAGCCAG GTCCTGGCAG**

251 **AGCAGCTCTG TCGACACATC TATGACTTGG CCTTGATGAC CCAAAAAGGG**

SalI

301 **TTTGATGCTG AGGGAATGAA AGCCTTCATT GAGCGTTCTA ATGCGGTCTT**

351 **GACGGCGTTG ACGACTCGCC AGTGA**AGGGT TTAGGGGAAC TCTAAACTTC

+1 T

401 ACAGAACGTC ATCCTAGCTA TCTCGCCCTC AGACGTGAGC CTCGCTAAGC

+1 P

451 TGAGGGCGAC G*ATCGCAGCG GAAAAATGTG ACCGGGGCGG CGCAG*ATC

Sau3A1 Sau3A1

Figure 1. Nucleotide sequence of the E10 promoter-active fragment.

The nucleotide sequence shown is the 3'end of the 1.75 kb E10 fragment immediately upstream of the promoter-less *GUS* gene. The 3'end of htpG heat shock protein gene (233 nucleotides) are in bold. The stop codons of htpG are marked by double lines. The true and processed transcription initiation sites of ORF76 are indicated by +1 T and +1 P, respectively (see Fig. 8A). The +1 P site is identical to the major transcription initiation site of E10-GUS determined using 5'RACE_PCR method. The putative LysR transcription activator binding site (CCTT-N_{11}-AAGG) and photoregulated motif (GG-N_5-AAC) are indicated by double dash line and dash line, respectively. The italic sequence of 34 bp between the Sau3A1 sites is an artifact from cloning.

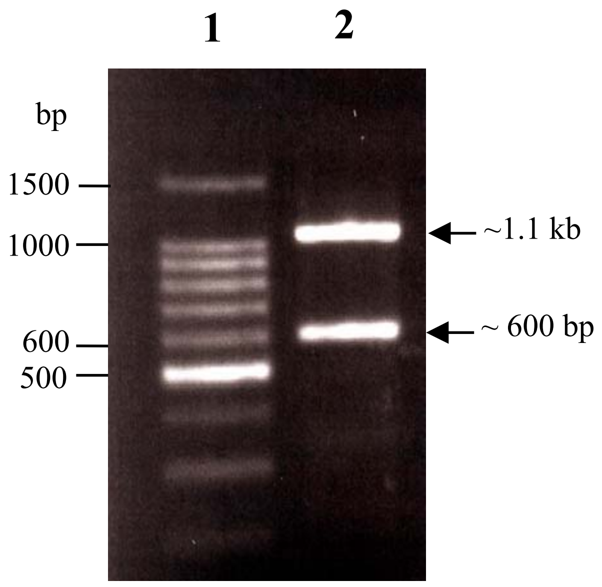

Figure 2. PCR products of downstream sequence of E10 fragment

1% agarose gel stained with ethidium bromide. Lanes 1, 100 bp marker; 2, PCR products of PvuII genomic walking library using primer Dwn-htpG1 and Adapter Primer 1.

A.

```
   1 GATCCGGCTA ATATACCCGA TGACCATGTC CTGCTCGTGA ATACCGCGCA
  51 TCCTCTTGTC CAGAACATCC TTAGTCTGCA GCAGGGGGCT ATCCTCAGCA
 100 GTGATGGACA CTCTCCAAGC CAGGTCCTGG CAGAGCAGCT CTGTCGACAC
 151 ATCTATGACT TGGCCTTGAT GACCCAAAAA GGGTTTGATG CTGAGGGAAT
 200 GAAAGCCTTC ATTGAGCGTT CTAATGCGGT CTTGACGGCG TTGACGACTC
 251 GCCAGTGAAG GGTTTAGGGG AACTCTAAAC TTCACAGAAC GTCATCCTAG
 300 CTATCTCGCC CTCAGACGTG AGCCTCGCTA AGCTGAGGGC GACGATCCAT
 351 TTCGTGGTGT GGGGTATGGG GTTGACTCTG TTTTCGCGAT CGCGGCGTCG
 400 CATCCCAAGC GGCCGTACCA ATATTATTCC GACGTCTCGA ACTTCAAGAA
 451 GAGCTTACCT ACGGGCGGGG ATGGCCTTGC TTCAGCTGAC TGCCTTAATT
 500 ATGGTTTTGG TGTGGTTGCG CGATACCCCC CGAGGGGCTT GGGCTGAGAG
 551 CGATTGCCAA AACTCTGTCT TTTCTAGTCT CTGTCGCGAT CGCTAATTTA
 601 TTCAATTATT TCTTCTGAAG AAGTATCTAA TAAAATAATG CGATCGCGCC
 651 CCTGCTGTTT GGCTTGATAC AGAGCACGAT CAGCACGTTC AAGGAATTGA
 701 TCAAGTGTCA GAATTTCAGG TGTCCAGACT GCAATGCCAA CACTAATGGT
 751 AATGGCGATC GCTTGTTGAG GAAGATGGAG AGATTGCTGA ATAACGAGCT
 801 GGCGGATGCG ATCCATCACT TGATATGCTT GCTTCCCATT GGTTTCAGGA
 851 AGCATTAATA CAAACTCATC CCCCCCCAGC CGAGCAAAAA TATCCACTTC
 901 GCGAATATTT TGTTGGCAAA TATTGGTCCA GAGAATTAAA ACCTGATCGC
 951 CAACCGTATG ACCGTAGCTG TCATTAATTG TTTTGAAATG ATCGATGTCA
1001 ATGATCACTA AAGAGAGTGA AGTGTCACGT TGGATACGAT AGATTCGGTT
1051 AAGTTCAGCT TGAGCAAGTG TAATAAACTG ACGGCGATTC AGAATGCCAG
1101 TCAGGGCATC TGTTGTTGCC TGTTGCTGCA ATTTTTGTTC TAGCTTCTTT
1151 TGTTCAGTAA TGTCAGTTGA AATACCGAGT AATCCAATAA TATTACTACT
1201 TTGTGCCGAT TCATAGATAG GCGTTTTGAT TTCTAGGTAG ACTCGTAATT
1251 GTCCATCTTT AGGGTTGACG AGATGGATTT CTTCCGTAGT CTGCTCTCCT
1301 TGGAAAACAC GAGAGTCGAT AGCTTGTAAT AATTTTGCAG CCTCTTCGGA
1351 GAAAAAGAAA AAATCATCCT TGCCCAAGAT TTCTTCCAAA GAGCGATCAA
1401 CAAAATCTAA AGCCGATCTA TTGGCATAGA TGTAACGAAA CTGAAGGTCT
1451 TTCATATAGA TATGAGAGCT GACCTGATCG AGTGCCGCTC GAAGACGTTG
1501 TGCTTCTACT AAAGCTTCTT GAAGTTGTTG TTTTGTTTTT CTCTCAG
```

3'end of *htpG*

ORF76

partial ORF

B.

```
            *      *        *  **
  1 MGLTLFSRSR RRIPSGRTNI IPTSRTSRRA YLRAGMALLQ LTALIMVLVW
       *
 51 LRDTPRGAWA ESDCQNSVWS SLCRDR
```

C.

```
  1 ERKTKQQLQE ALVEAQRLRA ALDQVSSHIY MKDLQFRYIY ANRSALDFVD
 51 RSLEEILGKD DFFFFSEEAA KLLQAIDSRV FQGEQTTEEI HLVNPKDGQL
101 RVYLEIKTPI YESAQSSNII GLLGISTDIT EQKKLEQKLQ QQATTDALTG
151 ILNRRQFITL AQAELNRIYR IQRDTSLSLV IIDIDHFKTI NDSYGHTVGD
201 QVLILWTNIC QQNIREVDIF ARLGGDEFVL MLPETNGKQA YQVMDRIRQL
251 VIQQSLHLPQ QAIAITISVG IAVWTPEILT LDQFLERADR ALYQAKQQGR
301 DRIILLDTSS EEIIE
```

D.

Figure 3. Downstream sequence of E10 promoter-active fragment.
A). Nucleotide sequences at downstream of the E10 fragment harbor the 3' end of *htpG* gene, an ORF of 76 amino acids (designated ORF76) and a partial ORF, B). The deduced amino acids of ORF76. The putative phosphorylation sites of protein kinase C are indicated by asterisks, C). The 315 deduced amino acids of the partial ORF. The sequence belongs to C-terminal of a hypothetical protein, D). Schematic representation of downstream sequence of E10 promoter-active fragment. The PAS and GGDEF/DUF1 domains are indicated. The figure is not drawn to scale.

A.

```
PCC7942: 16 QRLRAALDQVSSHIYMKDLQFRYIYANRSALDFVDRSLEEILGKDDFFFFSEEAAKLLQA 75
            +RLRA L+ +   +++ DL  R +YAN +A + +  S EE++GK        E  + LQ
    PAS:  1 ERLRAILESLPDGVFVLDLDGRILYANPAAEELLGYSPEELIGKSLLELIHPEDREELQE 60

PCC7942: 76 IDSRVFQ 82
               R+
    PAS: 61 RLQRLLS 67
```

B.

```
PCC7942: 143 ATTDALTGILNRRQFITLAQAELNRIYRIQRDTSLSLVIIDIDHFKTINDSYGHTVGDQV 202
             A  D LTG+ NRR F    + EL R  R Q  + L+L+++D+D+FK IND+YGH VGD+V
  GGDEF:   8 AAHDPLTGLPNRRYFEEELEQELQRARRQQ--SPLALLLLDLDNFKRINDTYGHAVGDEV  65

PCC7942: 203 LILWTNICQQNIREVDIFARLGGDEFVLMLPETNGKQAYQVMDRIRQLVIQQSLHLPQQA 262
             L         ++R  D+ ARLGGDEF ++LP+T+ + A ++ +RIR+L+    +
  GGDEF:  66 LQEVAQRLSSSLRRSDLVARLGGDEFAILLPDTSLEGAQELAERIRRLLAALKIPHTLSG 125

PCC7942: 263 IA--ITISVGIAVWTPEI-LTLDQFLERADRALYQAKQQGRDR 302
             +   +TIS+GIA + P      +  L+RAD+ALYQAK QGR+R
  GGDEF: 126 LPLYVTISIGIAAYPPNDGEDPEDLLKRADQALYQAKNQGRNR 168
```

C.

```
PCC7942: 143 ATTDALTGILNRRQFITLAQAELNRIYRIQRDTSLSLVIIDIDHFKTINDSYGHTVGDQV 202
             A  D LTG+ NRR F    + EL R    ++ +  +L++ID+D+FK IND+YGH VGD++
   DUF1:  12 AFRDPLTGLPNRRLFEEELEQELQR--AQRQGSPFALLLIDLDNFKQINDTYGHAVGDEL 69

PCC7942: 203 LILWTNICQQNIREVDIFARLGGDEFVLMLPETNGKQAYQVMDRIRQLVIQQSLHLPQQA 262
             L          +R  D+ ARLGGDEF ++LPET+ ++A  + +RI Q  +++ + +
   DUF1:  70 LQEVAQRLSSCLRPGDLVARLGGDEFAILLPETSLEEAIALAERILQ-QLREPIIIHGIP 128

PCC7942: 263 IAITISVGIAVWTPEILTLDQFLERADRALYQAKQQGRDRIIL 305
             + +TIS+G+A W       +  L+ AD ALYQAK+ GR+++ +
   DUF1: 129 LYLTISIGVAAYPNPGEDAEDLLKLADTALYQAKKAGRNQVAV 171
```

Figure 4. Amino acid alignment of the sequence III.

The 315 deduced amino acids of sequence III downstream of the E10 promoter-active fragment (see Figure 3) was highly homologous to conserved domains PAS, GGDEF and DUF1. A). The PAS conserved domain with length of 67 residues (accession no. smart00091). There are 34% identity and 52% positive. B). The GGDEF conserved domain with length of 168 residues (accession no. PF00990). There are 43% identity and 63.1% positive. C). The DUF1 conserved domain with length of 172 residues (accession no. smart00267). There are 38% identity and 62.3% positive.

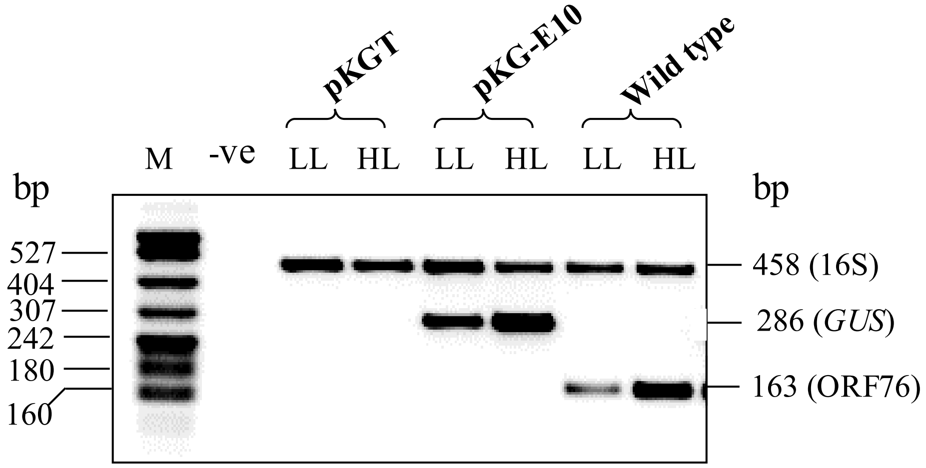

Figure 5. Comparison of the ORF76 and *E10-GUS* transcriptional levels using relative RT-PCR.

Negative image of 2% agarose gel containing ethidium bromide. *Synechococcus* was grown under low light intensity (LL, 1,500 lux) or high light intensity (HL, 5,000 lux). The cDNA of ORF76, *GUS* and 16S were synthesized and used as templates in the PCR reactions.The ORF76 and *GUS* cDNA were amplified for 25 cycles, whereas cDNA of 16S rRNA, used as internal control, was amplified for 10 cycles.

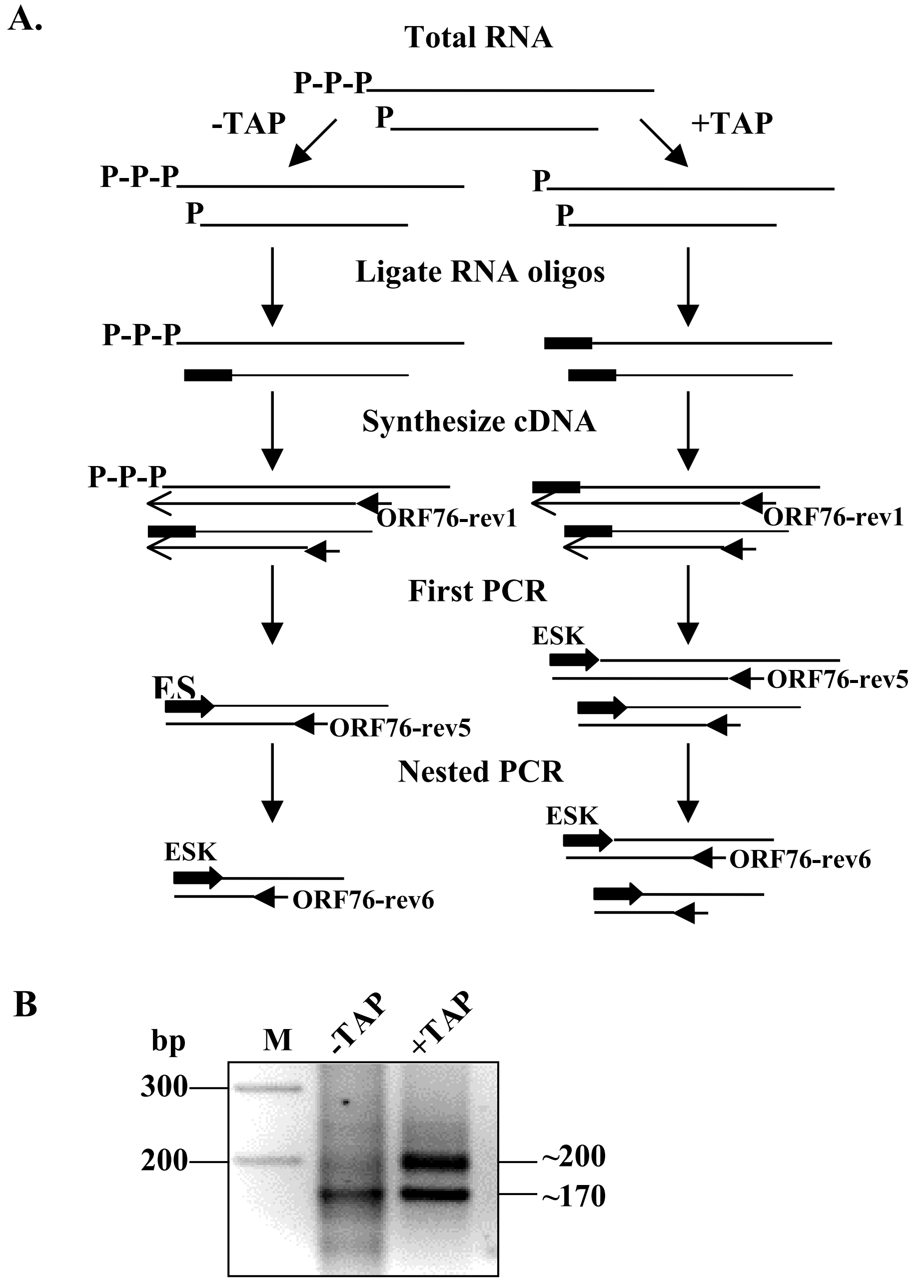

Figure 6. Analysis of the 5' end of *ORF76* transcripts.

A). Flowchart of the procedure used to differentiate true from processed transcripts of *ORF76*. Duplicate samples of total RNA were untreated or treated with TAP to convert 5' triphosphates to monophosphates. The 35-mer RNA oligos was ligated to 5' monophosphate end. The cDNA was synthesized using primer ORF76-rev1 and amplified using primers ESK and ORF76-rev6. Nested PCR was performed with primers ESK and ORF76-rev5. PCR products were separated by agarose gel electrophoresis (see Fig. 6B). The PCR product appearing in TAP-treated sample but not in untreated sample would have had 5' triphosphates which indicates that the 35-mer RNA oligos tagged the true initiated transcripts.

B). Negative image of 3% agarose gel containing ethidium bromide. PCR products generated from *ORF76* cDNA. Lane M, DNA size marker. Nested PCR of cDNA derived TAP-treated and untreated RNA are indicated. The PCR products of size around 200 and 170 bp result from true initiated and processed transcripts, respectively.

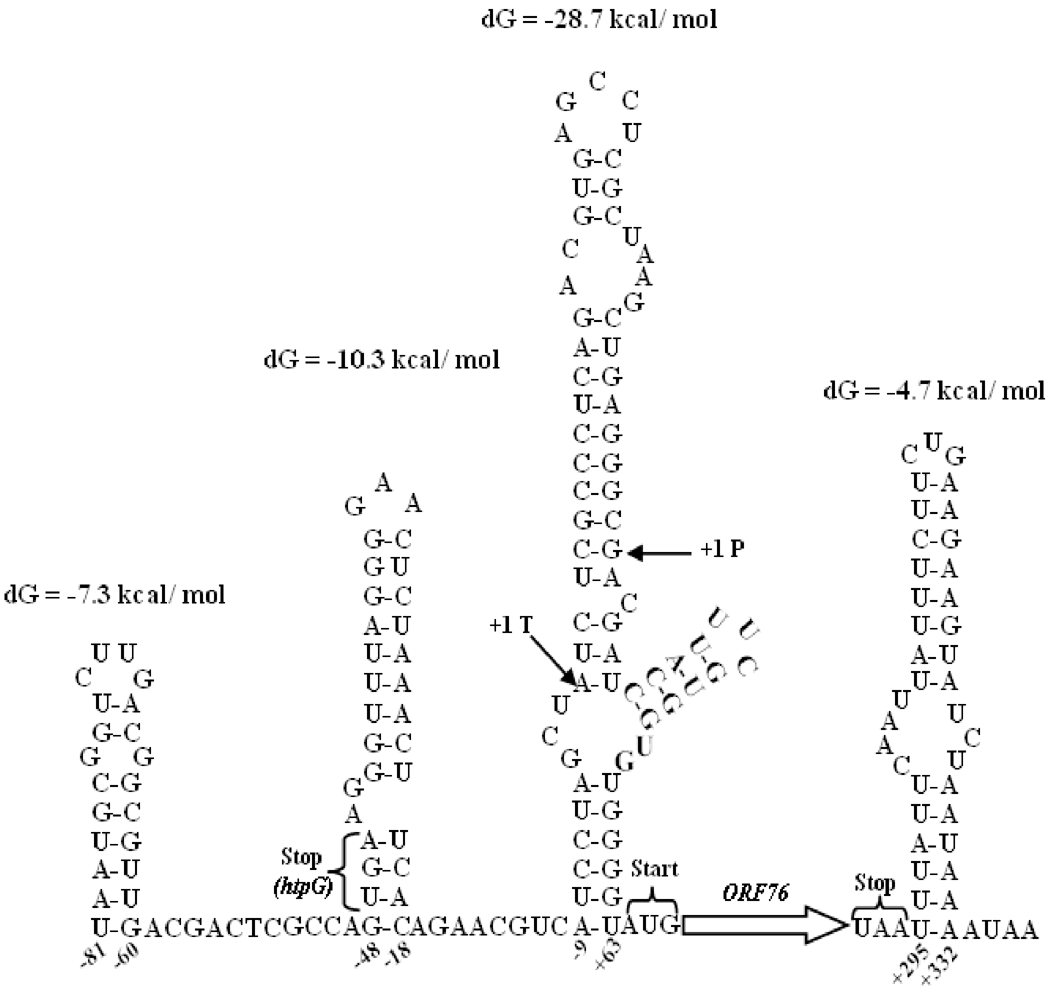

Figure 7. Predicted secondary structure of upstream and downstream regions of *ORF76* gene.
The upstream and downstream sequences of *ORF76* gene are written in the RNA form and their secondary structures are predicted. The delta G value are indicated. Positions of the true and processed transcription initiation sites are indicated by +1 T and +1 P, respectively. The nucleotides are indicated with respect to the +1 T site of *ORF76*.

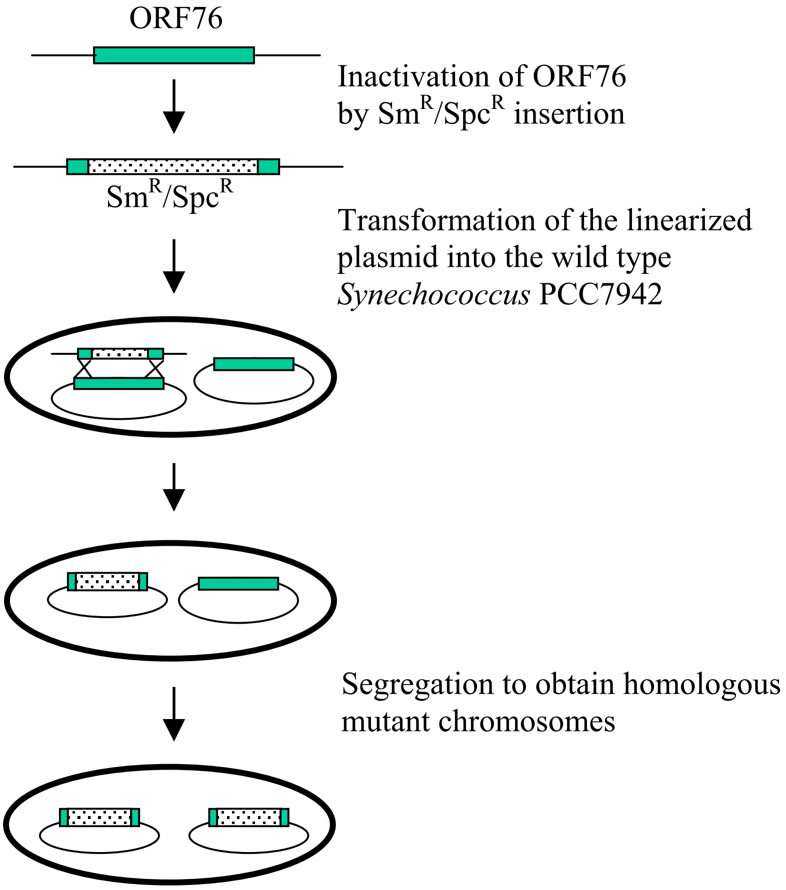

Figure 8. Flowchart of the procedure used to inactivate the *ORF76* in cyanobacterial chromosomes

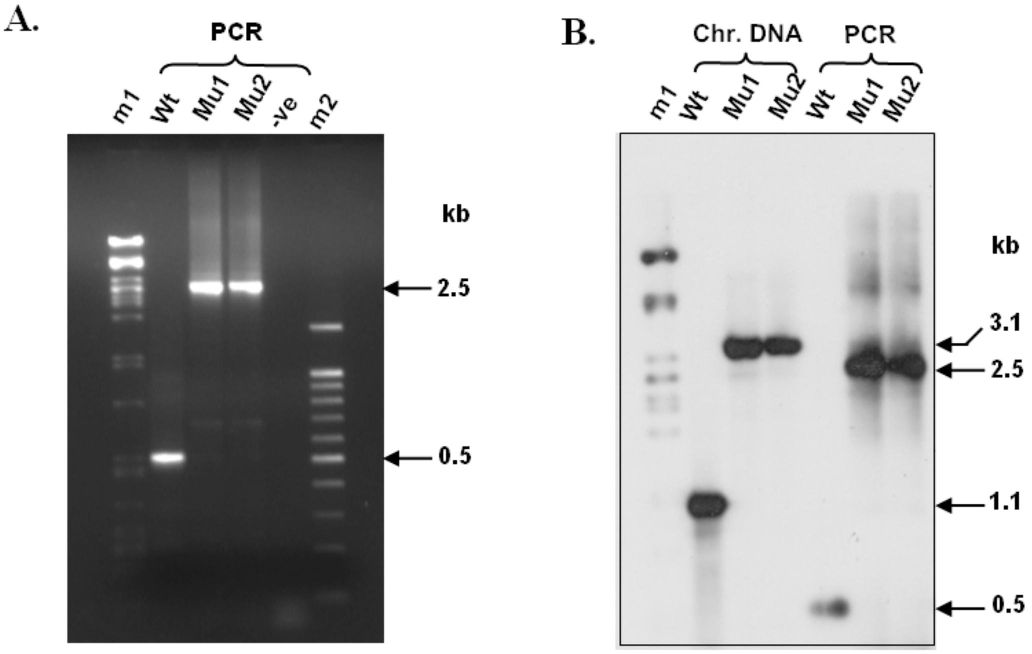

Figure 9. Southern blot analysis of ORF76 inactivated strains

A). 1% Agarose gel containing ethidium bromide.

B). Southern blot. Lanes m1, and m2 are DNA size markers, lambda DNA cut with PstI and 100 bp ladder, respectively. The PCR products were amplified with primers E10fok and ORF-rev3. -ve is negative control of PCR without DNA templates.Wt is of wild type *Synechococcus* PCC7942. Mu1 and Mu2 are of *ORF76* mutant strains.

RESISTANCE OF TRANSGENIC PAPAYA PLANTS TO PAPAYA RINGSPOT VIRUS THAI ISOLATE.

Sunee Kertbundit*, **Panu Ruangchan, Duriya Chantasingh, Gulsiri Charoensilp, Nuanwan Pongtanom, Orathai Tanwanchai, Pongsopee Attasart, Sakol Panyim and Mila Juricek**
Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakhon Pathom 73170, Thailand

Running title : PRSV coat protein gene transformation of Thai papaya

Abstract

The binary vector containing the cassette of the coat protein gene of *papaya ringspot virus*, the kanamycin resistance gene and the β-glucuronidase gene was constructed. This plasmid was used to transform the embryogenic calli and the somatic embryos of Thai papaya cultivar Khakdum by particle gun bombardment. After bombardment transgenic calli were selected on the media containing 100 mg/l kanamycin. Seven transgenic calli were obtained from 1200 clumps of somatic embryos bombarded with M10 tungsten particles (0.58%) and four transgenic calli from 780 clumps of somatic embryos (0.51%) were obtained from somatic embryos. PCR analysis of these transgenic plants showed that all contained CP gene as intact and truncated DNA fragment. Southern blot analysis revealed that several of transgenic lines contain more than one copy of CP gene inserts. Five transgenic lines showed resistance to PRSV infection. Detection of the transcriptional products of CP gene in these transgenic plants by RT-PCR technique showed that four resistance lines (G2, G3, G5 and T3) produced very low amount of CP gene transcription and no product was detected in PRSV resistant transgenic line T6. The CP protein product detection in these transgenic lines by western blot analysis showed the CP protein products detection only in transgenic line G1 and T2 which are not resistant to PRSV. The inheritance of the PRSV resistance to the progeny was analyzed in transgenic line G2 (R0). The resistance was transferred to the R1, R2 and R3 generation of transgenic line G2.

Keywords: papaya, papaya ringspot virus, coat protein, transgenic papaya

Introduction

Papaya (*Carica papaya* L.) is a fruit crop grown in tropical and subtropical countries. In Thailand, papaya is one of the staple food. The green fruits are used as salad and the ripened fruits are consumed as fresh dessert. Papaya trees are commercially grown in plantations and on a small scale in home gardens. Since 1975, the papaya production in Thailand has been limited by papaya ringspot virus (PRSV). The major cultivars of papaya grown in Thailand, Khak Dum and Khak Nuan are very susceptible for this virus. This virus causes seriously damage to papaya plantation. The infected area increases each year and it is now endemic throughout Thailand.

PRSV is a positive single stranded RNA virus in the potyvirus group (Purcifull *et al.* 1984) with the virions being flexuous, filamentous particles of 780x12 nm. PRSV is classified into two types according to the host range, type P, which infects papaya and some cucurbits and type W, which infects only cucurbits. These two types are not distinguishable by serological methods. PRSV type P can infect papaya at any growth stage and the infected papaya plants show a range of symptoms including yellowing and vein-clearing of younger leaves, mottling, leaves becoming severely distorted and narrow, ringed spots on the fruit as well as dark green streaks on the petioles and stems. Infected papaya plants exhibit significant stunting of growth as well as a reduction in both quality and quantity of fruit production. Several techniques have been applied to control the disease in the past including mild strain cross protection and conventional breeding programs, however, none of these techniques proved to be successful to date.

Coat protein mediated resistance (CPMR) has been successfully used to produce transgenic plants resistant to viruses (Baulcombe 1996, Beachy 1997). Transgenic papaya obtained by *Agrobacterium* and microprojectile bombardment had been reported (Fitch et al, 1990; Fitch et al, 1992, Yang et al, 1996, Cheng et al, 1996). In these experiments the transgenic Hawaiian papaya cultivar "Sunset" containing the coat protein gene of PRSV mild strain isolated from Hawaii (HA 5-1) were generated. This transgenic plant showed resistance to the severe strain of PRSV isolated from Hawaii but not to PRSV isolated from Thailand and other countries (Tennant et al. 1994). This result suggested that the resistance to the PRSV is strain specific and it must therefore be targeted to strains from the same geographical region. In our work, we constructed the binary vector containing the coat protein (CP) gene of PSRV Thai isolate (Ratchaburi) and transformed into Thai papaya cultivar Khak Dum using microprojectile bombardment. Eleven transgenic papaya lines containing the CP gene of PRSV Thai isolates had been generated in our laboratory. The resistance of transgenic papaya plants to papaya ringspot virus Thai isolate have been analysed..

Materials and Methods

Plasmid

The binary vector pSA1006 is based on pGV941 (Deblaere *et al.*, 1987) and contains the *nos* promoter- *nptII* gene-3'*nos* as a selectable marker and the CaMV35S promoter-*uidA* gene-3'*nos* from pBI 121 (Jefferson, 1987) as a reporter gene.

Plant material and culture.

Immature zygotic embryos were dissected from seeds of immature fruits of papaya var. Khakdum. Zygotic embryos were cultured on the induction medium (M1 medium) which consisted of: half strength MS salts medium (Murashige and Skoog's, 1962) plus 50 mg/l myo-inositol, full strength MS vitamins, 400 mg/l glutamine, 10 mg/l 2,4-D, 6% sucrose, 8 g/l Difco Bactoagar, pH 5.8. After the zygotic embryos were maintained in the dark at 26ºC for 3-4 weeks, the embryogenic calli which produced one or more somatic embryos on the apex were developed. These embryogenic calli were used in this transformation experiment. The second type of papaya culture used in this experiment is the somatic embryos. A small clump of somatic embryos was developed on the embryogenic callus when the callus was maintained in the induction medium for two to three months with monthly subcultured. Both types of tissues were transferred to the fresh M1 medium in group of 30 calli per petri dish one week before the bombardment

Plant transformation.

Embryogenic calli and somatic embryos were bombarded with M10 tungsten and 1.0 μm gold particles coated with the binary vector using the PDS 1000/helium-driven biolistic device (Bio-Rad, Hercules, Calif.). The gap distance between the rupture disk and the macrocarrier is 1.2 cm and the petri dish was placed at the target level 2 (6.0 cm). The chamber is evacuated to 26 in Hg and the helium gas pressure is set at 1300 psi. After bombardment the embryogenic calli and somatic embryos were cultured in the M1 medium for 2 days and transferred to the callus propagation medium (M1 medium supplemented with 75 mg/l kanamycin and 500 mg/l cefotaxime). After 14 days the tissues were transferred to the selective media which is the callus propagation medium containing 100 mg/l kanamycin. The tissues were subcultured monthly to fresh selective media until resistant embryogenic calli developed.

***Plant regeneration*.**

The resistant embryogenic calli were transferred to glass vessels with plastic lids containing the regeneration medium (MS medium supplemented with 0.2 mg/l BAP, 0.1 mg/l kinetin, 3% sucrose, 8 g/l Difco Bacto agar, 100 mg/l kanamycin, pH5.7). The tissues were cultured under 12 hour photoperiod using cool white Sylvania fluorescent lamps until the embryos regenerated and formed shoots. Shoots with 3-4 leaves and about 1.5 cm tall were sliced and transferred to the rooting medium contains the full strength MS salts, full strength MS vitamins, 100 mg/l myo-inositol, 0.05 mg/l indolbutyric acid (IBA), 3% sucrose, 8 g/l Difco Bactoagar at pH 5.8. Rooting plants were acclimatized by replaced the plastic lids of the culture vessels with a sheet of double layer sterile tissue paper and covered with saran wrap for 2 weeks before the plant transferred to 1:1 mixture of soil and vermiculite.

PRSV inoculation.

The PRSV isolate Ratchaburi was propagated in normal papaya plants. The infected leaves were ground in 0.01 M phosphate buffer, pH 7.0. The leaf extract was rubbed onto the two upper leaves celite-dusted of transgenic plants. After 10 minutes the inoculated leaves were rinsed with water. Symptoms of PRSV infection were observed for at least 90 days.

PCR analysis.

The coat protein gene insertion in the plants was detected by PCR. The total genomic DNA isolated from transgenic leaves of CTAB method (Rogers. and Bendich., 1994) was used as DNA template. The 50 µl of total PCR reaction was composed of 100 ng DNA template, 50 pmol of each forward and reverse primers, 200 µM of dNTPs mix, 1x PCR buffer (20 mM Tris-HCl pH 8.8, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton x-100, and 100 µg/ml nuclease-free BSA), 0.5 µl (1.5 units) *Pfu* DNA polymerase and water. The amplification was carried out in the gene amplification system 2400 thermal cycle (Perkin-Elmer) under the following condition : 94 °C for 30 s, 50 °C for 30s, 72 °C for 2 min, for 35 cycles. The PCR product was analyzed in 0.8% agarose gel electrophoresis.

Southern blot hybridization.

…..en microgram of total DNA isolated from each transgenic plants was digested with the restriction enzymes *Hind* III. The resulting fragments were subjected to electrophoresis in 0.8% agarose gel, transferred to Hybond-N membrane (Amersham) and hybridized using the 810 bp coat protein amplified fragment as a probe. The probe was labeled with Gene Images Random-Prime Labeling and Detection System (Amersham) according to the manufacturer's protocol.

RT-PCR analysis.

Total RNA was isolated from transgenic leaves of each plants by TRIzol reagent kit (GIBCO BRL). The RNA was treated with DNaseI to remove the DNA. The first strand cDNA was synthesized in the reverse transcription reaction using enzyme Superscript ^{TM} II (GIBCO BRL). The 20 µl reverse transcription reaction reaction mixture contained 10 pmol of specific primers (5'CP and 3′CP), 2 µl of 100 ng purified viral RNA and sterile DEPC-treated distilled water. The reverse transcription reaction was performed using Gene Amplification System 2400 (Perkin Elmer). The reaction mixture was heated to 70°C for 10 min and quick chilled on ice. The 7 µl of 5x buffers (250 mM Tris-HCl pH 8.3, 375 mM KCl and 15 mM $MgCl_2$), 0.5 mM dNTPs and 0.01 M DTT were added to the reaction mixture and incubated at 42°C for 2 min. The 1 µl (200 units) of Superscript^{TM} II was added and the reaction mixture was incubated at 42°C for 20s and 50°C for 50 min. The reaction was inactivated at 70°C for 15 min and incubated at 4°C for 5 min. The 10-20 µl of this first strand cDNA/RNA hybrid was directly used as the template for PCR reaction.

Western blot hybridization.

Total proteins were isolated from calli and leaves of plants by TCA method (Wu and Wang, 1984). Thirty five microgram of each proteins were separated on a 13 % SDS polyacrylamide gel. The proteins were transferred to nitrocellulose membrane and reacted with a 1:2,000 dilution of the anti-CP polyclonal antibodies, and followed by incubation with a 1:5,000 dilution of the anti-mouse IgG conjugated with horseradish peroxidase. The signal was detected using the ECL plus western blotting detection reagent (Amersham Pharmacia Biotech).

Inheritance of the coat protein and resistance.

Transgenic papaya G2 which is a female papaya (R0) was cross-pollinated with the the hermaprodite Khak Dum papaya to set fruits. The seeds were collected from the fruits and sow on the soil to obtain the R1 plants. The hermaprodite R1 plants (G2-1-4) were self-pollinated to set fruits and the seeds from the fruits were grown to obtained the R2 plants. The one month old of these progeny plants were used for PRSV inoculation and molecular analysis.

Results

Transformation of Thai papaya cutivar Khakdum

The somatic embryos of Khak Dum papaya was bombarded with the binary vector pSA1006 (Fig.1). After bombardment the cultures were transferred to the selectable medium containing 100 mg/l kanamycin and transferred monthly to fresh selective medium until transgenic calli were selected which took about 4-5 months. Transgenic calli were identified as light-yellow somatic embryo clusters among the brown tissues of the non-transformed calli. After the tissues were cultured on selective medium for 5 months, seven transgenic calli were obtained from 1200 clumps of somatic embryos bombarded with M10 tungsten particles (0.58%) and five transgenic calli were obtained from 780 clumps of somatic embryos bombarded with 1.0 μm gold particles (0.51%).

Plant regeneration

The transgenic calli started to regenerate after culturing for 3-4 months on the regeneration medium containing NAA, BAP and kanamycin and multiple shoots were formed within 6-9 months. Shoots with 3-4 leaves and about 1.5 cm tall were sliced and transferred to the rooting medium.

PCR analysis of transgenic plants.

The Coat protein gene insertion in the transgenic lines were determined by PCR analysis. All of the transgenic lines showed 1 Kb amplification fragment of the coat protein gene except lines G3, G5 and T3 of which the truncated 0.8 Kb fragment of the coat protein gene were amplified (Fig 2). In transgenic line G2 both 1 kb intact and 0.8 kb truncated fragments of CP gene were presented.

Southern blot analysis.

The coat protein gene insertion in transgenic lines G1, G2, G3, G5, T1, T2, T3,T4, T5, T6 and T7 was analyzed by Southern blot analysis. The total plant genomic DNA was digested with restriction enzyme *Nco* I flanking the 1 kb coat protein gene and hybridized with the CP gene probe. The result show that most of the transgenic plants contains multiple copies except G1 and T2 (Fig. 3). There is no detected DNA band in transgenic line T1.

RT-PCR analysis in transgenic plants.

The transcriptional products in 11 transgenic lines were determined by RT-PCR analysis. All of the transgenic lines showed 1 Kb amplification fragment of the coat protein gene except lines G2, G3, G5 and T3 of which the truncated 0.8 Kb fragment of the coat protein gene were amplified (Fig 4). The intensity of the amplified bands of the coat protein gene in transgenic lines G2, G3, G5, T3 and T5 is lower than other transgenic lines. No RT-PCR amplification product was detected in transgenic line T6.

Western blot analysis in transgenic plants.

The protein product of the inserted CP gene in eleven transgenic plants was analysed by western blot analysis. The CP protein band of 32 Kda was detected only in transgenic lines G1 and T2 (Fig 5. and 6.).

PCR and physical analysis of the coat protein gene insertion in progeny plants.

The inheritance of the PRSV resistance was further analyzed in G2 line. All G2 plants were grown to maturity in a greenhouse. They had a normal morphology and growth habit compared to the non transgenic control. Segregation of the introduced genetical material was investigated by PCR and Southern blot analysis on 13 plants. The coat protein gene insertion in the 13 progeny R1 plants of transgenic line G2 was analyzed by PCR. The truncated coat protein gene fragment of about 0.8 kb was detected in 9 plants (G2-1-5-2, G2-1-5-8, G2-1-6-1, G2-1-4-2, G2-1-3-1, G2-1-3-4, G2-1-2and G2-1-4) but not in 4 plants (G2-1-2-7, G2-1-1, G2-1-3 and G2-1-6) (Fig 7. and 8.). The coat protein gene insertion in these plants was confirmed by Southern blot analysis. The total plant genomic DNA from these transgenic lines was digested with restriction enzyme *Nco* I flanking the 1 kb coat protein gene. Hybridization of plant DNA digested with restriction enzyme *Nco* I showed that the progeny plants containing the CP gene detected by PCR showed multiple copies of the coat protein gene inserted pattern as in transgenic line G2 (Fig 9. and 10). The PCR analysis of the 42 progeny R2 plants showed that 78% of the progeny R2 plants contain the truncated coat protein gene fragment of about 0.8 kb (Fig. 11, 12 and 13). Southern blot hybridization of the plant genomic DNA from these R2 progeny plants digested with restriction enzyme *Nco* I showed multiple copies of the coat protein gene inserted pattern as in transgenic line G2 and the R1 progeny plants (Fig. 14 and 15).

Resistance to PRSV infection in progeny plants of transgenic line G2.

Sixty progeny plants from transgenic G2 crossed with normal papaya were tested for PRSV infection under greenhouse condition. Thirty progeny plants showed complete resistance.

Forty two progeny plants from R2 progeny (self pollinated line G2-1-4 plants) were tested for PRSV infection under greenhouse condition. Thirty three progeny plants (78%) showed complete resistance. There is no symptom on the inoculated leaves nor on new emerging leaves during plant maturation over than 6 months.

Discussion

The efficiency of transformation of papaya cultivar. Khakdum using somatic embryos is comparable to the previously reported in particle gun bombardment of zygotic embryos of papaya cultivars "Sunset" (Fitch *et al*, 1990) and " Maradol" (Cabrera-Ponce *et al*, 1995). However the papaya cutivar Khak Dum took about 1 month longer to regenerate. The genotype of the papaya, the type and age of tissue used in the particle gun bombardment are the important factors that effect the transformation efficiency by particle gun bombardment .

Plant transformation by microprojectile bombardment results in transgene integration patterns that generally exhibit multiple transgene copies and extensive rearrangements of the introduced DNA (Pawlowski and Sommers, 1996). It has been shown originally on rice, that transgene loci in plants transformed by particle gun bombardment have host DNA separating closely linked transgene sequences (Kohli *et al*, 1998). Fig. 3 shows that similar situation was observed in papaya transformation. Only two transgenic papaya lines have a single copy gene (G1 and T2), while there are three distinct patterns showing rearranged transgenes. This is probably due to the fact that direct DNA transfer is mediated by two-phase integration mechanisms and it is using integrational hot spots (Kohli *et al*., 1998).

However, only lines G2, G3, G5 and T3 with the same integration pattern are resistant to the virus. The majority of virus resistance in plants has been shown to occur on RNA level. The underlying mechanism of virus resistance in plants is believed to be post transcriptional gene silencing (PTGS) (Wang and Waterhouse, 2002). As a result of PTGS, RNA accumulation is suppressed, which can be seen from Fig. 2A, where no RNA was amplified by RT-PCR from virus resistant lines. The induction of PTGS by transgene is believed to be facilitated by creation of inverted repeats separated by a short spacer (Waterhouse, 2001). Such a specific rearrangement did not take place in transgenic lines T4, T5, T6 and T7, where no resistance to PRSV was found.

The PCR amplified fragment from lines resistant to PRSV shows a missing 166 bp from the 3’ end of the CP gene as determined by sequencing (data not shown). Although it has been reported that the truncated CP is able to induce resistance against potyvirus (Silva Rosales *et al*., 1994, Leclerc and AbouHaidar, 1995), it seems unlikely in this case. No above mentioned putative inverted repeat was found in this 0.8 kb fragment. Also, the 166 bp fragment missing from the 0.8 kb fragment can be found in all other upper bands in Fig.1.

The protein involvement in this resistance was not found - only two lines with intact expression cassette (G1, T2) show protein expression. The ORF and/or the promoter are apparently rearranged in the non resistant lines as well. Only TEV resistance was shown to be mediated by protein so far (Marcos and Beachy,1997).

The PCR analysis and Southern blot analysis of R1 and R2 generations and analysis of segregation patterns showed integration of the target genes to have occurred at one or few loci, resulting in stable and predictable pattern of inheritance, which shall prevail through future generations. Such integration in a single or few loci could not have taken place by a chance alone, however, the mechanisms governing this phenomenon are not yet known.

Acknowledgements

We would like to thank Surapee and Kittisak Kiratiya-ankul for providing us the PRSV samples. This work was funded by Mahidol University Research grant and Thai Research Fund (TRF) grant.

References

1. Baulcombe DC (1996) Mechanisms of Pathogen-derived Resistance to Viruses in Transgenic Plants. Plant Cell 8:1833-1844
2. Beachy RN (1997) Mechanisms and applications of pathogen-derived resistance in transgenic plants. Curr Opin Biotechnol 8:215-220
3. Cabrera-Ponce JL, Vegas-Garcia A, Herrera-Estrella L (1995) Herbicide resistant transgenic papaya plants produced by an efficient particle bombardment transformation method. Plant Cell Rep 15:1-7
4. Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1990) Stable transformation of papaya via microprojectile bombardment. Plant Cell Rep 9:189-194
5. Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1992) Virus Resistant Papaya Plants Derived from Tissues Bombarded with the Coat Protein Gene of Papaya Ringspot Virus. Bio/Technology 10 :1466-1472
6. Rogers S.O. and Bendich A.J., 1994. Extraction of total cellular DNA from plants, algae and fungi. Plant Molecular Biology Manual D1:1-8
7. Leclerc D, AbouHaidar MG (1995) Transgenic tobacco plants expressing a truncated form of the PAMV capsid protein (CP) gene show CP-mediated resistance to potato aucuba mosaic virus. Mol Plant Microbe Interact 8:58-65
8. Murashige T, Skoog, F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473-497
9. Purcifull DE, Edwardson JR, Hiebert E, Gonsalves D (1984) Papaya ringspot virus. CMI/AAB Description of Plant Viruses 292
10. Silva-Rosales L, Lindbo JA, Dougherty WG (1994) Analysis of transgenic tobacco plants expressing a truncated form of a potyvirus coat protein nucleotide sequence. Plant Mol Biol 24:929-939
11. Srisomchai T (1975) Studies on papaya ringspot virus, NEROA (Norteast Regional Office of Agricalture) Report. 228-232
12. Tennant PF, Gonsalves C, Ling KS, Fitch MMM, Manshardt RM, Slightom JL, Gonsalves D (1994) Differential Protection Against Papaya Ringspot Virus Isolates in Coat Protein Gene Transgenic Papaya and Classically Cross Protected Papaya. Phytopathol 84:1359-1366
13. Yang J-S, Yu, T-A, Cheng, Y-H, Yeh S.-D. (1996) Transgenic papaya plants from *Agrobacterium*-mediated transformation of petioles of *in vitro* propagated multishoots. Plant Cell Rep 15:459-464
14. Pawlowski WP and Somers DA. (1996) Transgenic DNA integrated into the oat genome is frequently interspersed by host DNA. Proc Natl Acad Sci 13;95(21):12106-10.
15. Kohli A, Leech M, Vain P, Laurie DA and Christou P. (1998) Transgene organization in rice engineered through direct DNA transfer supports a two-phase integration mechanism mediated by the establishment of integration hot spots. Proc Natl Acad Sci U S A. 1998 9;95(12):7203-8.
16. Wang MB and Waterhouse PM (2002) Application of gene silencing in plants. Curr Opin Plant Biol.;5(2):146-50.
17. Marcos JF and Beachy RN. .(1997)Transgenic accumulation of two plant virus coat proteins on a single self-processing polypeptide J Gen Virol.;78 (Pt 7):1771-8.

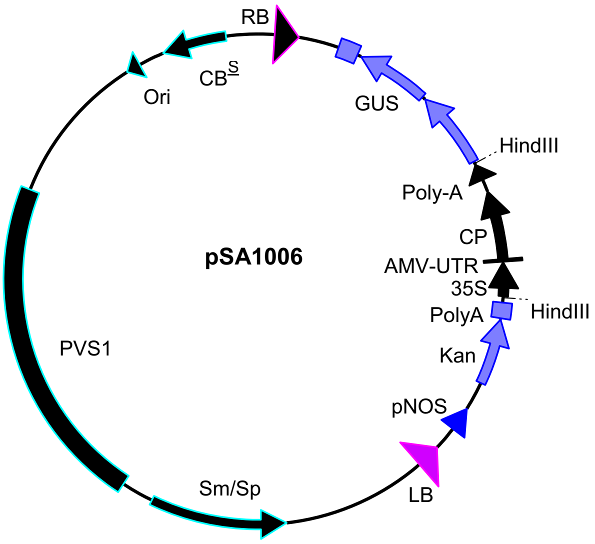

Figure 1. Schematic map of binary vector pSA1006. The CaMV 35 S promoter-*gus* coding sequence-3'*nos* is at the right border

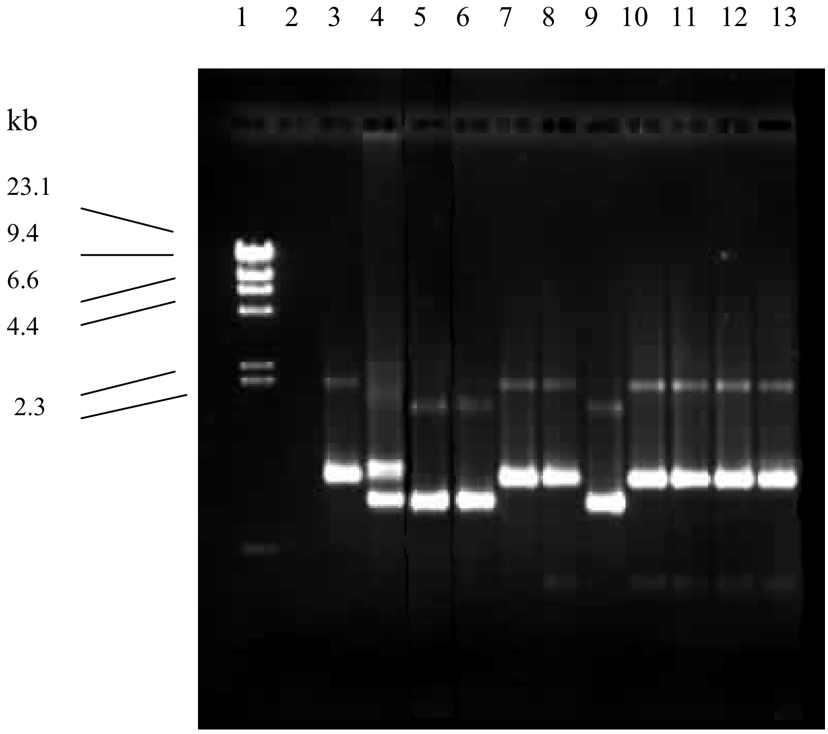

Figure 2. PCR amplification of the coat protein gene inserted in the transgenic lines. The PCR products were analyses in 0.8% agarose gel, stained with ethidium bromide. Lane1; λ-DNA digested with *Hind* III as a marker; Lane 2, PCR product of non-transgenic papaya plant, Lane3 to 7; PCR products of transgenic lines G1, G2, G3, G5, T1, T2, T3, T4, T5, T6 and T7 respectively

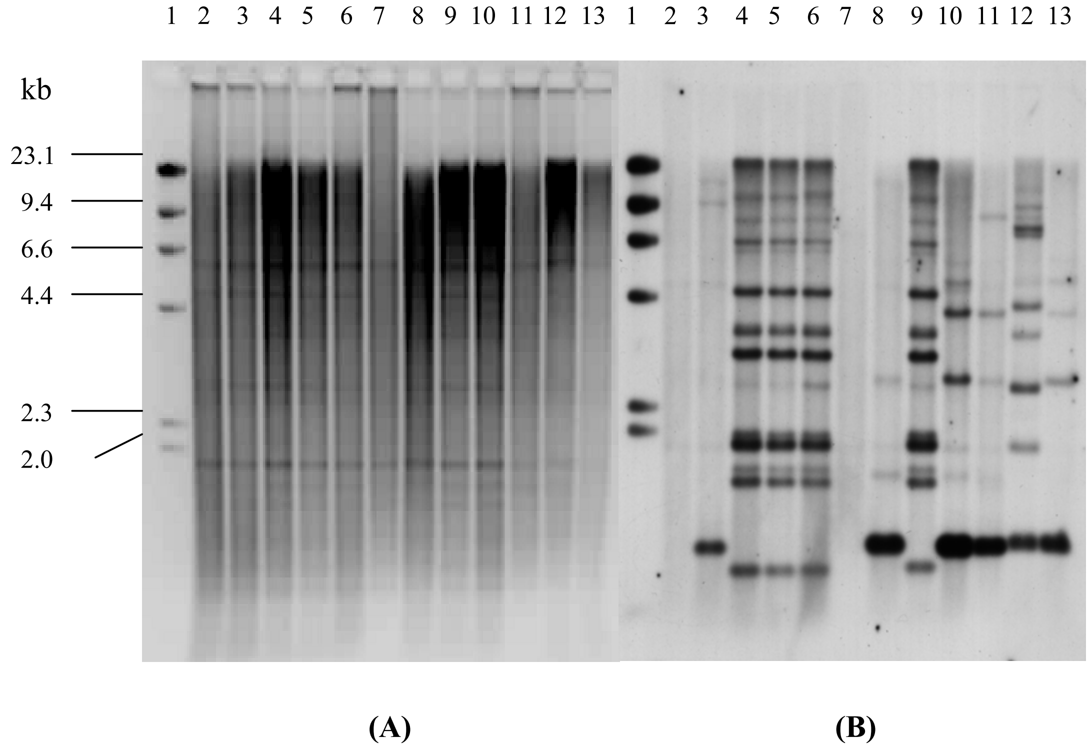

Figure 3. Agarose gel electrophoresis (A) and Southern blot hybridization (B) of transgenic lines. Total genomic DNA from each plant was digested with restriction enzyme *Nco* I which flanks the PRSV coat protein. The DNA was analysed in 0.8% agarose gel, transferred to nylon membrane and hybridized with DNA probe. Lane 1: λ DNA digested with *Hind* III, Lane2: DNA of non transgenic papaya, Lane 3-13: DNA of transgenic line G1, G2, G3, G5, T1, T2, T3,T4, T5, T6 and T7.

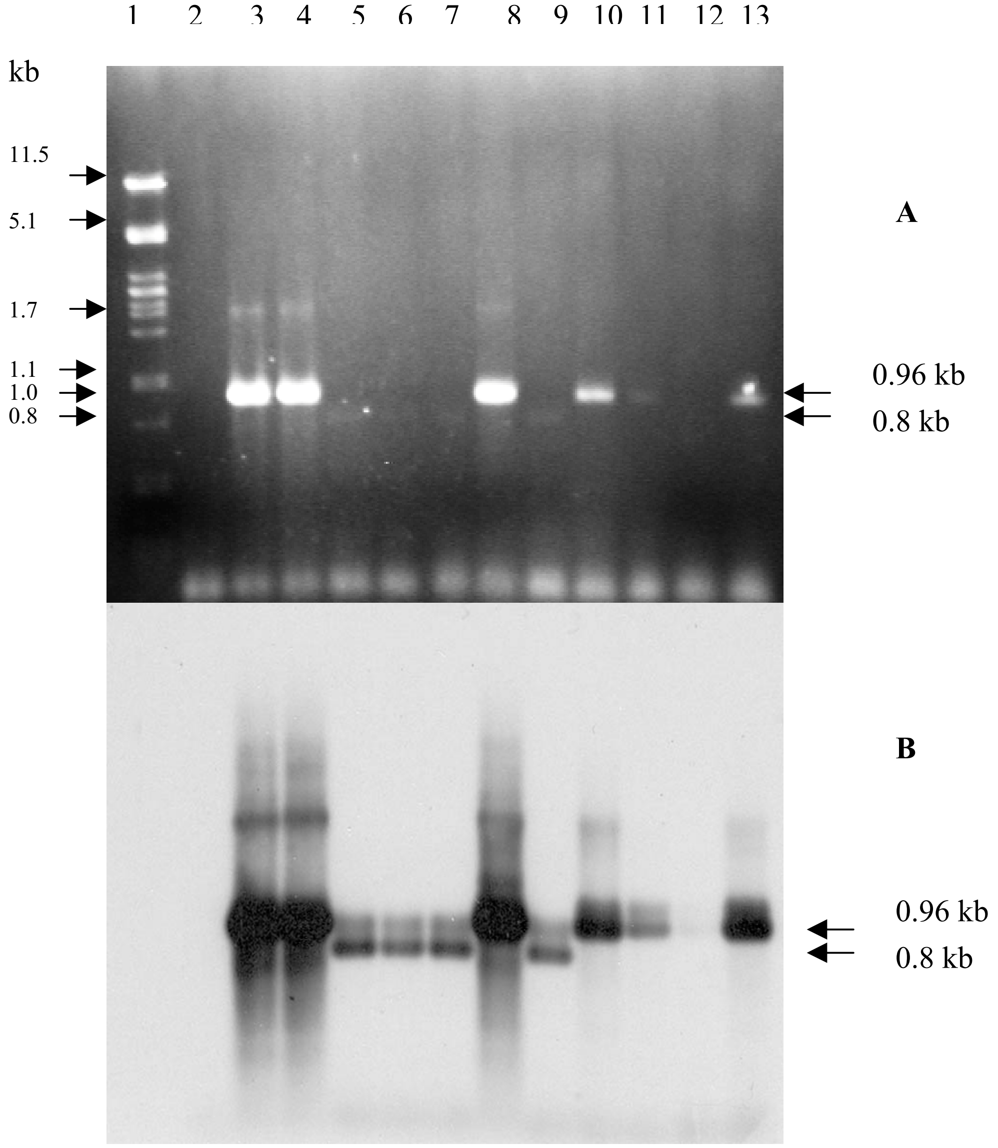

Figure 4. RT-PCR amplification of the coat protein gene inserted in the transgenic lines. The PCR products were analysed in 0.8% agarose gel (A) transferred to nylon membrane and hybridized with probe(B). Lane 1: λ DNA digested with *Pst* I, Lane2: PCR amplification without DNA template (negative control), Lane 3: PCR amplification using pSA1006 as a DNA template (positive control), Lane 5-13: PCR product of transgenic lines G2, G3, G5, T2, T3, T4, T5, T6 and T7.

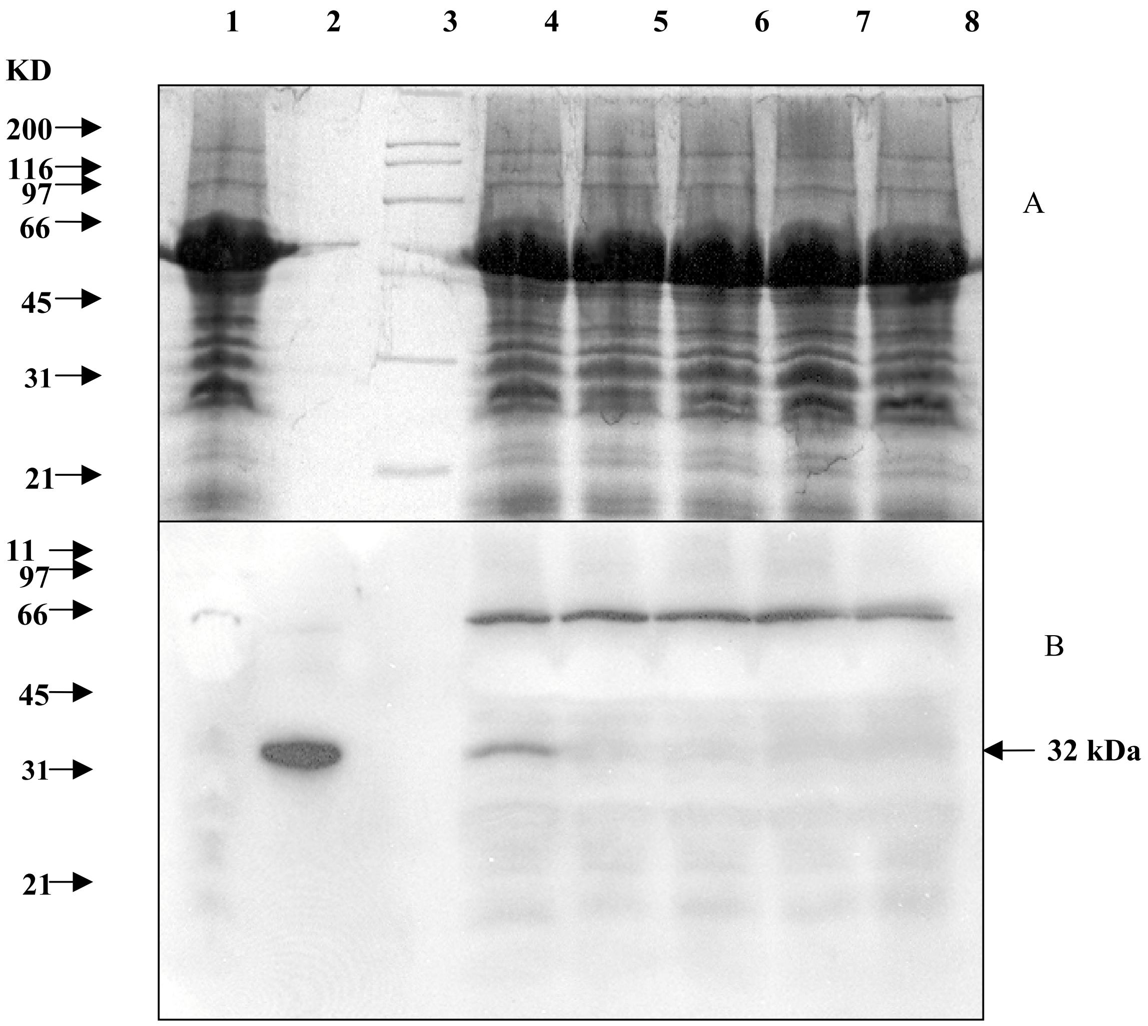

Figure 5. SDS-PAGE (A) and western blot analysis (B) for detection of CP protein in the transgenic papayas. Total proteins were isolated from the transgenic papayas, healthy papaya and infected papaya and thirty five microgram from each proteins were separated on a 13 % SDS polyacrylamide gel. The proteins were transferred to nitrocellulose membrane and were reacted with a 1:2,000 dilution of the primary antibody, anti-CP polyclonal antibodies, and followed by incubation with a 1:5,000 dilution of the secondary antibody. Lane 1: 35 μg of total protein extract from healthy papaya leaves, Lane 2: 500 ng of total protein extract from PRSV infected papaya leaves, Lane 3: Board range protein marker. Lane 4-8: 35 μg of total protein extract from transgenic papaya leaves line G1, G2, G3, G5 and T3 respectively.

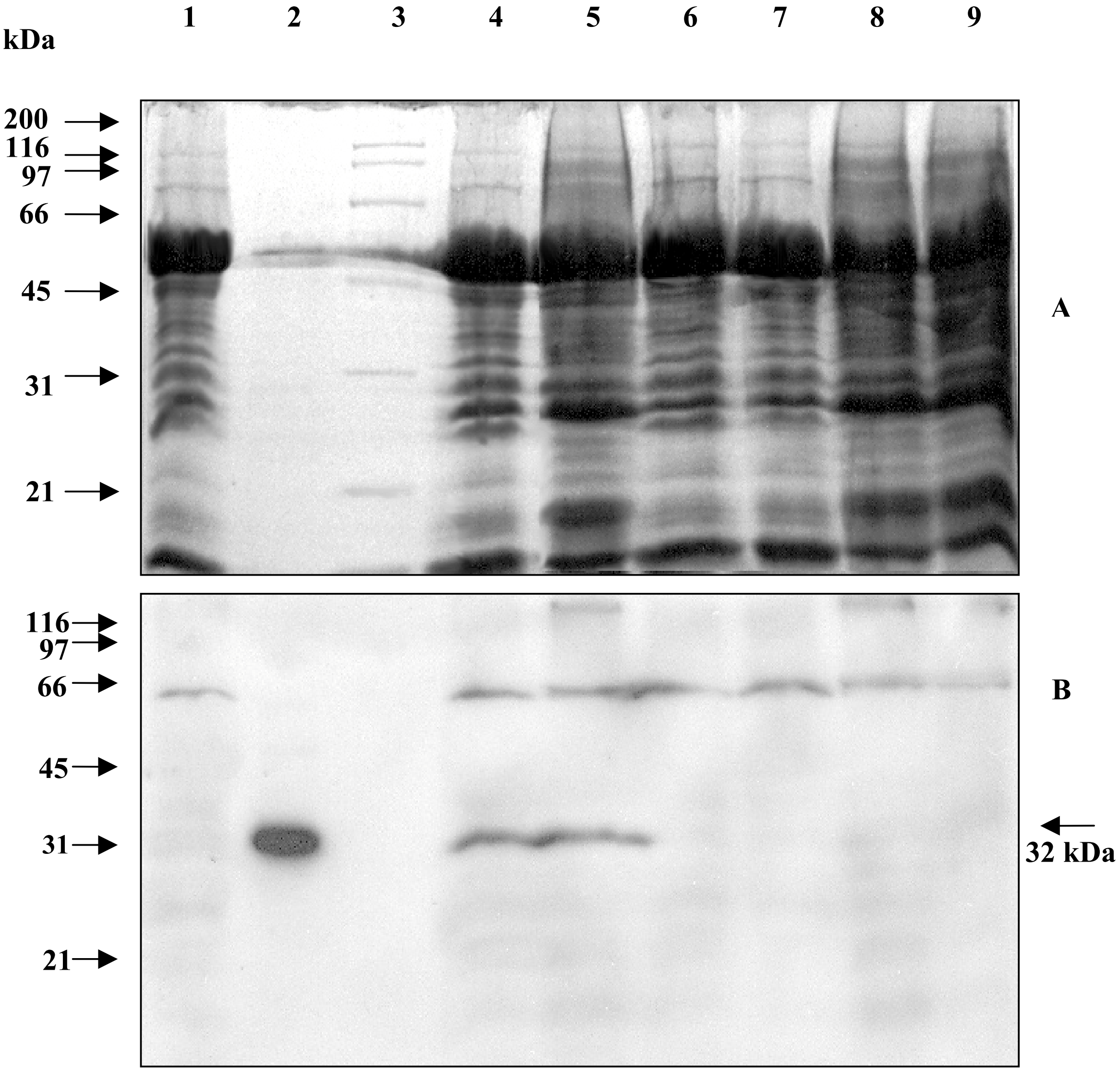

Figure 6. SDS-PAGE (A) and western blot analysis (B) for detection of CP protein in the transgenic papayas. Total proteins were isolated from the transgenic papayas, healthy papaya and infected papaya and thirty five microgram from each proteins were separated on a 13 % SDS polyacrylamide gel. The proteins were transferred to nitrocellulose membrane and were reacted with a 1:2,000 dilution of the primary antibody, anti-CP polyclonal antibodies, and followed by incubation with a 1:5,000 dilution of the secondary antibody. Lane 1: 35 µg of total protein extract from healthy papaya leaves, Lane 2: 500 ng of total protein extract from PRSV infected papaya leaves, Lane 3: Board range protein marker, Lane 4-9: 35 g of total protein extract from transgenic papaya leaves, G1, T2, T4, T5, T6 and T7 respectively

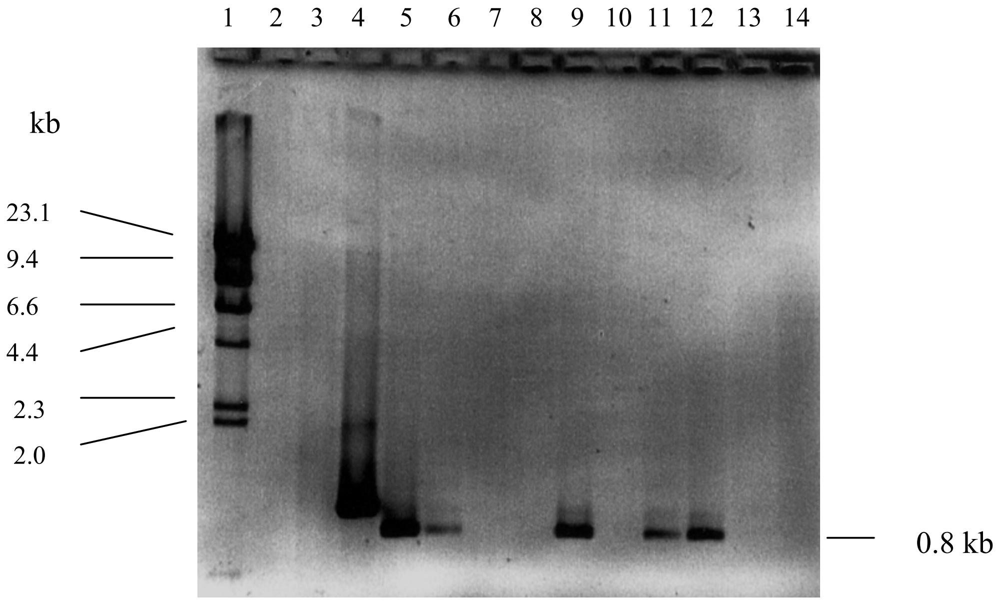

Figure 7. PCR amplification of the coat protein gene inserted in the R1 progeny of transgenic line G2. The PCR products were analysed in 0.8% agarose gel, stained with ethidium bromide. Lane 1: λ DNA digested with *Hind* III, Lane2: PCR amplification without DNA template (negative control), Lane 3: PCR product of non infected normal papaya (negative control), Lane 4: PCR amplification using pSA1006 as a DNA template (positive control), Lane 5: PCR product of transgenic line G2, Lane 6-13: PCR product of progeny plants G2-1-3-4, G2-1-2-7, G2-1-1, G2-1-2, G2-1-3, G2-1-4, G2-1-5 and G2-1-6, Lane 14: PCR product of non infected normal papaya (negative control).

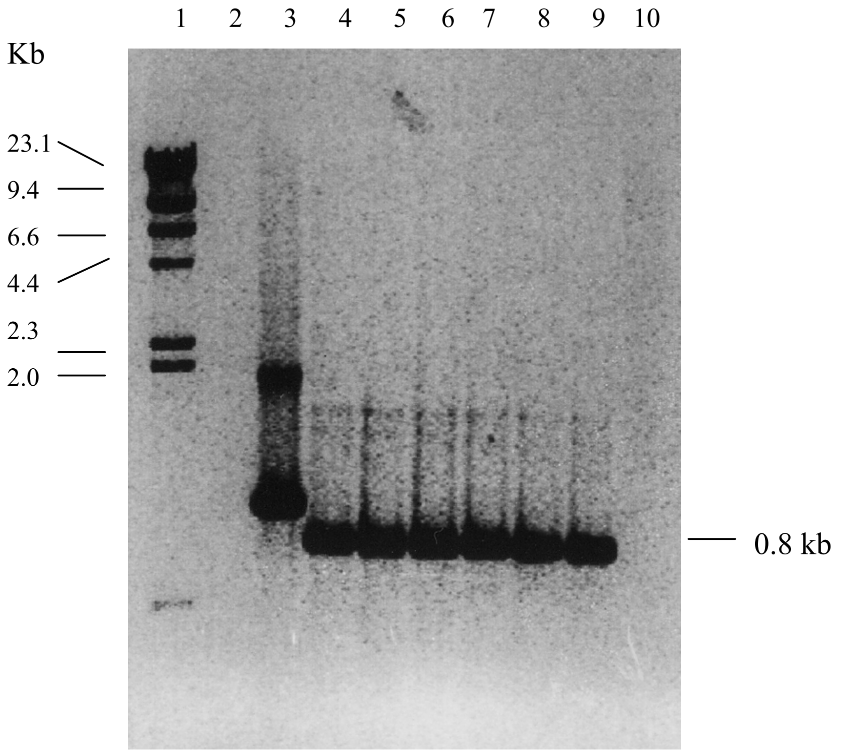

Figure 8. PCR amplification of the coat protein gene inserted in the R1 progeny of transgenic line G2. The PCR products were analysed in 0.8% agarose gel, stained with ethidium bromide. Lane 1: λ DNA digested with *Hind* III, Lane2: PCR amplification without DNA template (negative control), Lane 3: PCR amplification using pSA1006 as a DNA template (positive control), Lane 4-10: PCR product of progeny plants G2-1-5-2, G2-1-5-8, G2-1-6-1, G2-1-4-2, G2-1-3-1, G2-1-3-4 and G2-1-2-7.

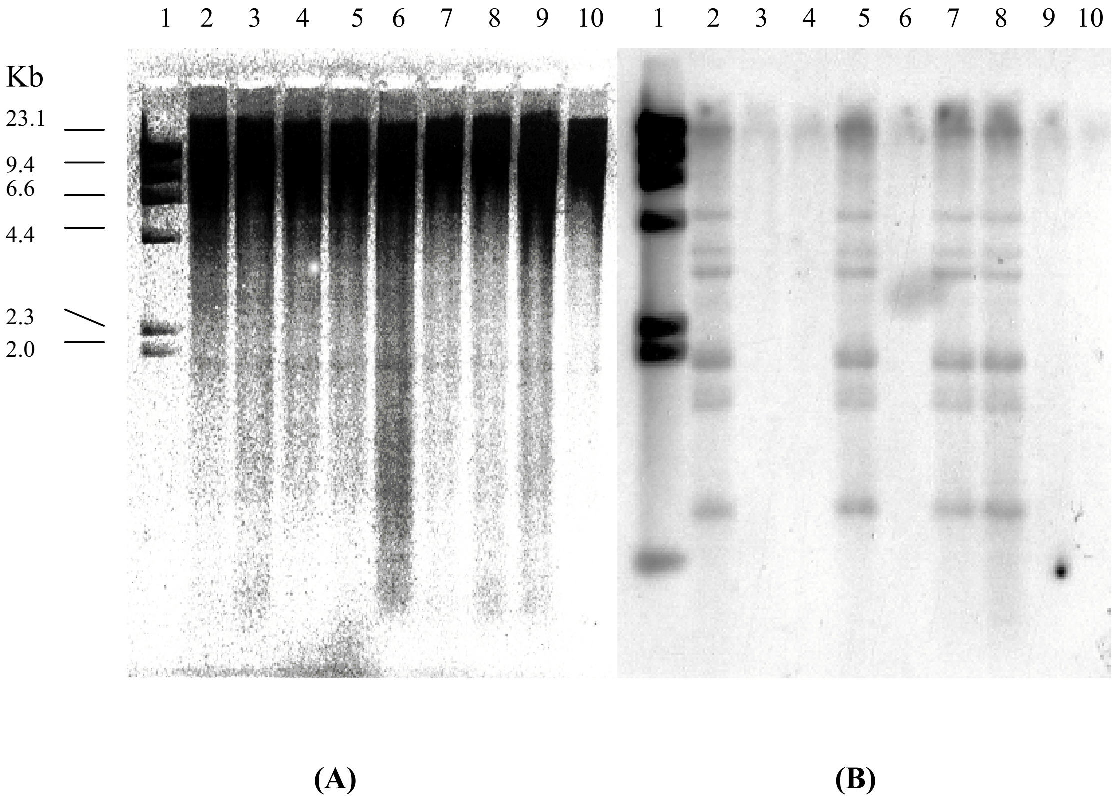

Figure 9. Agarose gel electrophoresis (A) and Southern blot hybridization (B) of R1 progeny plants of transgenic line G2. Total genomic DNA from each plants was digested with restriction enzyme *Nco* I which flanks the PRSV coat protein. The DNA was analysed in 0.8% agarose gel, transferred to nylon membrane and hybridized with DNA probe. Lane 1: λ DNA digested with *Hind* III, Lane2: DNA of transgenic line G2, Lane 3-9: DNA of progeny plants G2-1-2-7, G2-1-1, G2-1-2, G2-1-3, G2-1-4, G2-1-5 and G2-1-6, Lane 10: DNA of non transgenic papaya.

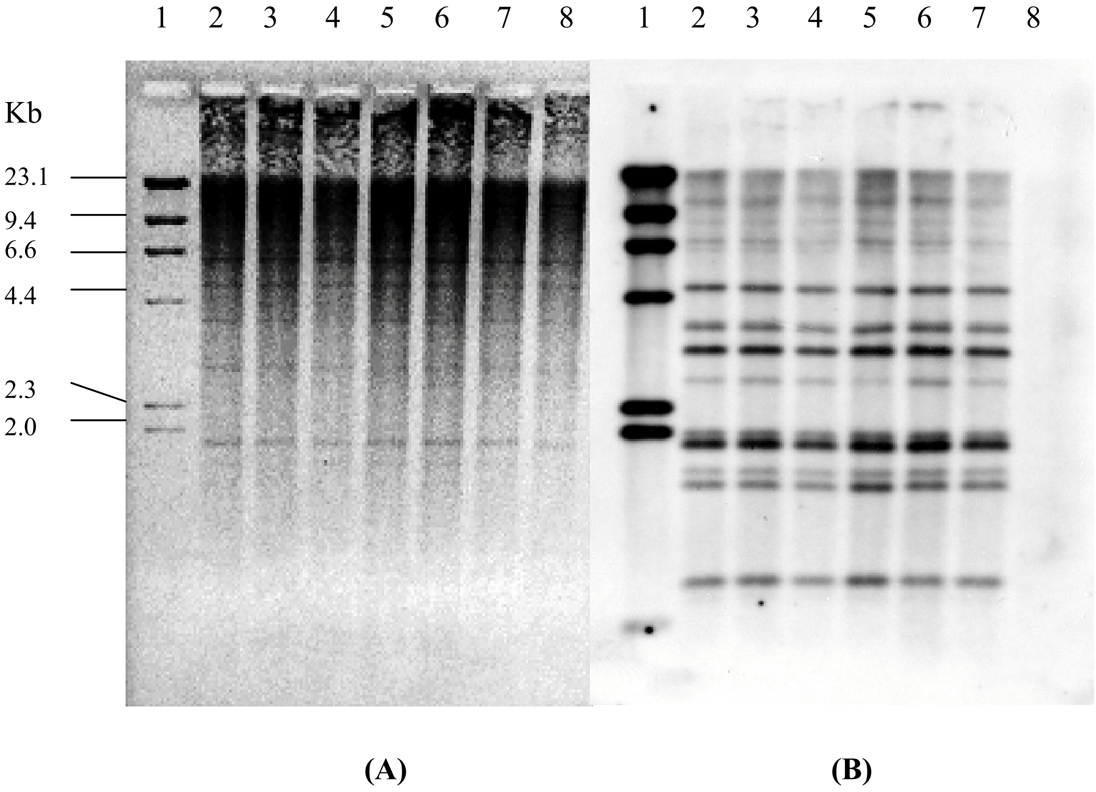

Figure 10. Agarose gel electrophoresis (A) and Southern blot hybridization (B) of R1 progeny plants of transgenic line G2. Total genomic DNA from each plants was digested with restriction enzyme *Nco* I which flanks the PRSV coat protein. The DNA was analysed in 0.8% agarose gel, transferred to nylon membrane and hybridized with DNA probe. Lane 1: λ DNA digested with *Hind* III, Lane2-8: DNA of progeny plants G2-1-5-2, G2-1-5-8, G2-1-6-1, G2-1-4-2, G2-1-3-1, G2-1-3-4 and G2-1-2-7.