

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Sequence analysis of rice *rubi3* promoter gene expression cassettes for improved transgene expression

Elumalai Sivamani^{a,*}, Joshua D. Starmer^b, Rongda Qu^{a,**}

^a Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7620, USA

^b Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

ARTICLE INFO

Article history:

Received 9 April 2009

Received in revised form 13 August 2009

Accepted 14 August 2009

Available online 21 August 2009

Keywords:

Exon

GUS

IME

rRNA

rubi3 promoter

Translation

ABSTRACT

In a construct containing a *GUS* reporter gene driven by the 5' regulatory elements from *rubi3*, expression was enhanced 4-fold when a 20-nucleotide (nt) *GUS* 5' untranslated sequence was replaced with 9 nt sequences derived from *rubi3*'s second exon. The roles of the sequences immediately upstream from the *GUS* translation initiation codon, and their significance in gene expression, were investigated. Sequence analysis suggests that complementarity between sequences immediately 5' of a translation initiation codon and the rice 17S rRNA may be responsible for the reduction in protein levels from constructs containing the *GUS* leader sequence. The results demonstrate an affect sequences immediately upstream from transgenic coding sequences have on expression, and when using the *rubi3* 5' regulatory sequence in particular.

© 2009 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Ubiquitin is a highly conserved eukaryotic protein that contains 76 amino acids (aa) and is important for many cellular processes [1]. The polyubiquitin gene family consists of genes containing tandem repeats of the 76 aa monomer [1,2]. In plants, many polyubiquitin genes have been isolated and studied [1,3]. Polyubiquitin gene promoters usually have strong and constitutive, or near-constitutive, expression patterns [4] and thus have a potential to be widely used in transgene studies in plants, as well as in crop biotechnology. In monocot transgenic research, the maize *Ubi1* promoter [6] has been widely used to express reporter genes and other “useful” genes [5]. The high expression capacity of these promoters is usually attributed to an approximately 1 kb sequence

immediately 5' of the translational start codon that is spliced out and not present in the mRNA [3,5] (referred to as a “5' UTR intron”).

We recently isolated and characterized the promoter region of a rice polyubiquitin gene, *rubi3*, that contained an 808 bp promoter, a 67 bp 5' UTR (referred to as “untranslated exon 1”, or “exon 1”), and an 1140 bp 5' UTR intron [3]. We observed over 20-fold intron-mediated enhancement (IME) of the *GUS* reporter gene expression in transient assays in rice suspension cells. Modifications of nucleotide sequences in exon 2, which contains the translation start codon and the first ubiquitin monomer coding sequence, further increased expression to 90-fold over the intronless *rubi3* promoter [3]. This independent, 4-fold enhancement took place at the translational level [7].

Many strategies have been used to augment transgene expression in plants. Among them, certain leader sequences in front of the AUG translation initiation codon increase efficiency of translation initiation to enhance transgene expression [8,9]. The best-known instance is the use of the omega enhancer sequence from the tobacco mosaic virus [10]. Some plant gene sequences also enhance transgene translation. For example, a 5' UTR sequence conserved among soybean lipoxygenase genes, when fused with a β -glucuronidase (*GUS*) gene under control of the CaMV 35S promoter, increased *GUS* activity by 11- and 7-fold in rice and tobacco, respectively [11]. A similar sequence from the pea plastocyanin gene, *petE*, stimulated 5–10-fold higher *GUS* translation in transgenic tobacco [12]. In addition, translational fusion of a

Abbreviations: aa, amino acid; GUS, β -glucuronidase; IME, intron-mediated enhancement of gene expression; LUC, luciferase; MU, methylumbelliferone; MUG, 4-methylumbelliferyl- β -D-glucuronide; NOS, nopaline synthase; nt, nucleotide; UTR, untranslated region; X-Gluc, 5-bromo-4-chloro-3-indolyl- β -D-glucuronidase.

* Corresponding author. Present address: Syngenta Biotechnology, Inc., PO Box 12257, 3054 E. Cornwallis Road, Research Triangle Park, NC 27709-2257, USA. Tel.: +1 919 281 2193/515 7616; fax: +1 919 541 8585/515 7959.

** Corresponding author. Tel.: +1 919 515 7616; fax: +1 919 515 7959.

E-mail addresses: sivamani.elumalai@syngenta.com (E. Sivamani), rongda_qu@ncsu.edu (R. Qu).

ubiquitin monomer coding sequence with a reporter gene also increased transgene expression [13,14]. In our earlier report, we demonstrated enhanced *GUS* expression in rice suspension cells by translationally fusing the *rub13* ubiquitin monomer coding sequence (from exon 2) with the *GUS* coding sequence in place of the 20 nt *GUS* non-coding “leader sequence” [3]. The enhancing effect caused by the ubiquitin monomer fusion was retained when the monomer sequence was narrowed down to 9 nucleotides, coding for the first three amino acid residues of the *rub13* protein. Mutagenesis at the third nucleotide of the AUG codon in this 9 nucleotide sequence still maintained the enhancing effect while producing the original *GUS* protein without any fusion [3]. This observation prompted us to investigate the roles of the 9 nt from *rub13* exon 2 and the 20 nt non-coding leader sequence in *GUS* expression. This 20 nt leader sequence consists of 14 nt from the original *GUS* gene and a 6 nt *SmaI* site coming from vector pAHC25 [3,6].

In this current study, we reduced the 9 nt from *rub13* exon 2 to 3 and 0 nt in the gene construct in order to determine the minimum sequence required to maintain enhanced *GUS* expression. We show that the leader sequence carried over with the *GUS* coding sequence largely suppresses this enhancement. We further demonstrate that removal of a similar leader sequence from a *GUS* construct under control of the maize *Ubi1* promoter yields an approximately 4-fold increase in *GUS* gene expression. Computational analysis of the 20 nt *GUS* leader sequence reveals substantial complementarity with the rice 17S rRNA, and this could be hindering translation.

2. Materials and methods

2.1. Plasmid constructs

A series of PCRs were performed to make gene constructs for the experiments. **Supplementary Table 1** lists the primers used for the PCRs. A 368 bp PCR product was amplified using FusionF and McsR2 primers with pRESQ4 [3] plasmid DNA as the template. The PCR product was then cloned in PCR4 TOPO™ (Invitrogen,

Carlsbad, CA, USA). The 272 bp DNA fragment obtained by digesting the resulting plasmid with *Bgl*III and *Sma*I was ligated to pRESQ4, which had been digested with the same two enzymes, to obtain pRESQ2b (Fig. 1).

A 373 bp PCR product obtained by using FusionF and 9aa* PCR primers and pRESQ36 [3] plasmid DNA as the template was cloned in PCR4 TOPO™. The 166 bp fragment from *Eco*RV and *Bs*MI enzyme digestions of the resulting plasmid was ligated in *Eco*RV/*Bs*MI digested pRESQ36 to get pRESQ85 (Fig. 2).

The 373 and 370 bp PCR products using primers FusionF and 3ntatcR or Resq4 with pRESQ48 [3] plasmid DNA as the template were cloned in PCR4 TOPO™. Digestion of the two intermediate plasmids with *Bgl*III and *Bs*MI resulted in 321 and 318 bp fragments, and these were ligated with pRESQ48 digested with the same enzymes to obtain pRESQ68 and pRESQ70 plasmids, respectively (Fig. 3). To make binary vectors for stable transformation, the 4136 or 4127 bp fragments obtained by digesting pRESQ48 or pRESQ70 with *Eco*RI enzyme were cloned in *Eco*RI-linearized pCAMBIA1300 (CAMBIA, Canberra, Australia) binary vectors to obtain pCAM48 and pCAM70, respectively.

A three-primer PCR technique [15] was used to make the maize *Ubi1-GUS* fusion gene construct, pRESQ95, based on pRESQ29 [3], to eliminate the 37 nt *GUS* leader sequence. The two outer primers, *Ubi*zlf and FusionR, were used in combination with the linking primer, *Ubi*fusion, to make a 1394 bp PCR product. This PCR product was cloned in PCR4 TOPO™. The 1008 bp fragment from the *Nco*I and *Bs*MI digestion of the resulting plasmid was ligated into *Nco*I/*Bs*MI digested pRESQ29 to obtain pRESQ95, which contained the maize *Ubi1* promoter and 5' UTR intron directly fused to the first ATG of the *GUS* coding sequence (Fig. 6).

A 433 bp PCR product obtained by using FusionR and Zerosma PCR primers and pRESQ70 plasmid DNA as template was cloned in PCR4 TOPO™. The 394 bp fragment from *Xma*I and *Sna*BI digestion of the intermediate plasmid was ligated into *Xma*I/*Sna*BI digested pRESQ4 to obtain pRESQ101 (Fig. 7). All the PCR products were subjected to sequence analysis before being used for gene construction.

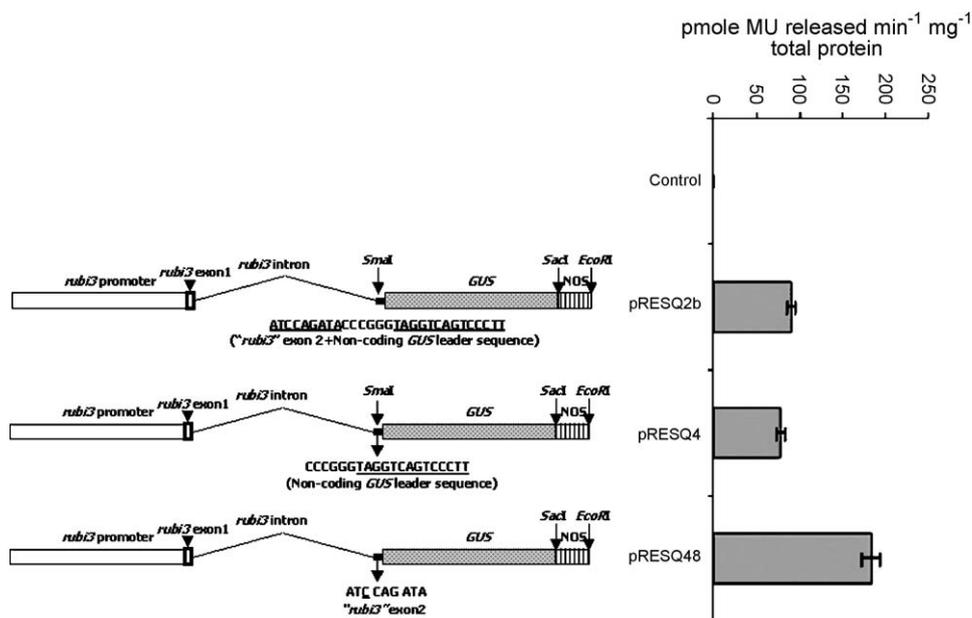


Fig. 1. The effects of the modified 9 nt *rub13* exon 2 sequences when placed along with the *GUS* non-coding leader sequence in the *rub13* promoter constructs. Plasmid construct pRESQ2b was made to have both leader sequences (underlined) from pRESQ48 and pRESQ4. pRESQ4, with a 20 nt *GUS* leader sequence, and pRESQ48, containing the modified 9 nt *rub13* exon 2 sequence, were used for comparison. The graph on the right shows *GUS* activity. Rice suspension cells bombarded with the luciferase gene (*LUC*)-containing plasmid pJD313 were used as a control, and the data in the graph was normalized to the LUC activity.

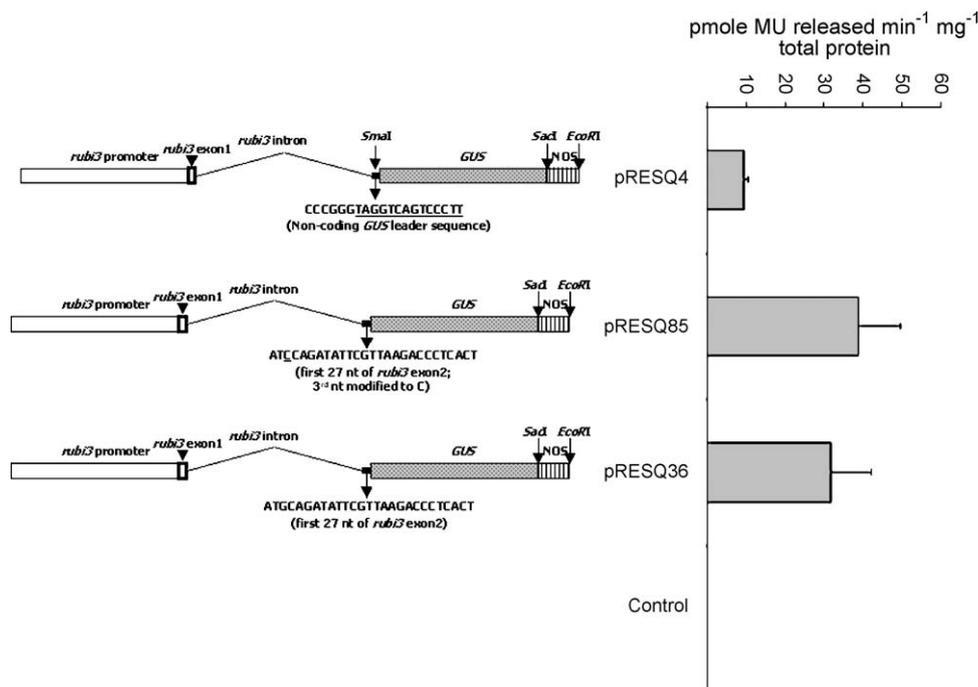


Fig. 2. The effects of distance between the 5' UTR intron and the *GUS* coding sequence in *rubi3* promoter constructs. Construct pRESQ36 fused the first 9 aa (27 nt) from *rubi3*'s exon 2 to the *GUS* protein and had enhanced expression over pRESQ4. Construct pRESQ85 contains the same 27 nt as pRESQ36, except a single nt mutation at the ubiquitin ATG start codon (underlined). pRESQ4 had 20 nt from the *GUS* leader sequence (underlined) between the intron and the *GUS* ATG start codon. Transient MUG assays were performed with these constructs using rice suspension cells. The data in the graph are normalized to LUC data. The control cells were bombarded with pJD313 only.

2.2. Transient *GUS* assays

All of the plasmids used for transformation experiments were purified using the Qiagen[®] plasmid midi kit (Qiagen, Valencia, CA, USA) according to the manufacturer's procedure. Particle bombardment of various *GUS* gene constructs on rice suspension cells was employed to evaluate the expression levels of the constructs. Culture conditions, bombardment methods, and transient *GUS* assays were performed using procedures described previously [3]. In all transient assays, each construct was co-transformed with pJD313 containing a luciferase gene (*LUC*) expression cassette to serve as an internal control for normalizing the *GUS* data. Controls were maintained in every experiment, in which suspension cell

were bombarded with gold particles coated with pJD313 plasmid only. All assays had three replicates for statistical analysis.

2.3. Stable transformation of rice with *Agrobacterium tumefaciens*

The freeze–thaw method [16] was applied in transformation of *Agrobacterium tumefaciens*. The binary vector DNA from pCAM48 and pCAM70 was used to transform the *A. tumefaciens* strain EHA105 [17] carrying supervirulent plasmid pTOK47 [18] to yield EHA105 (pTOK47, pCAM48) and EHA105 (pTOK47, pCAM70), respectively. Rice (*Oryza sativa* L. cv. Taipei 309) transformation was performed using EHA105 (pTOK47, pCAM48) or EHA105 (pTOK47, pCAM70) following the protocol published by CAMBIA

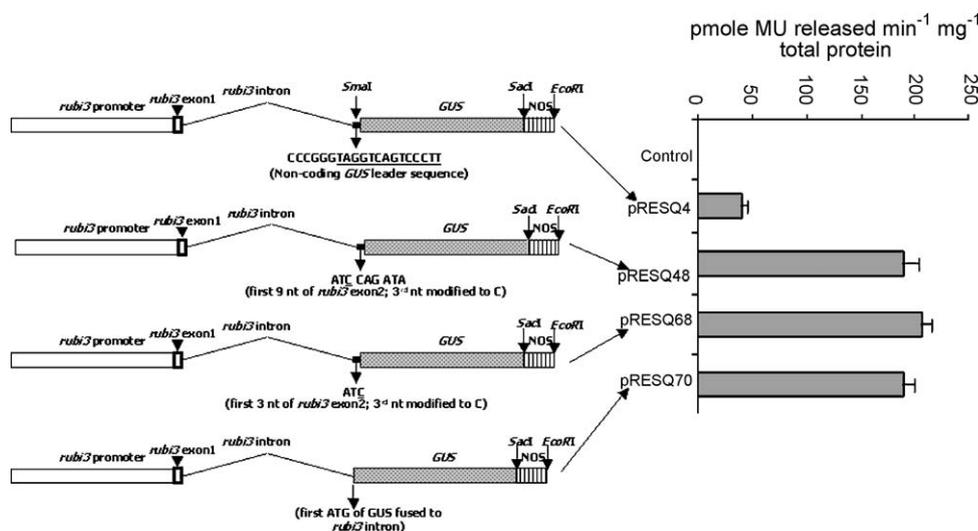


Fig. 3. Transient *GUS* assays to determine minimum requirement of *rubi3*'s exon 2 subsequence for expression enhancement. Plasmid constructs (pRESQ48, 68 and 70) containing various lengths of mutated subsequences from *rubi3*'s exon 2 were co-transformed with pJD313 containing the *LUC* gene in rice suspension cells and compared with pRESQ4. The graph shows data normalized to LUC activity in each of these constructs. The control cells were bombarded with pJD313 only.

[19]. Calluses induced from mature seeds were infected with *Agrobacterium* and selected with hygromycin B (AG Scientific Inc., San Diego, CA, USA). After rooting, the regenerated plants were transferred to pots and maintained in a greenhouse.

2.4. Northern blot analysis of the transgenic rice lines

Total RNA from the T1 seedlings of various transgenic rice lines was isolated using Trizol[®] reagent (Invitrogen) according to the manufacturer's procedure. The first leaf of an approximately 4-week-old seedling was used for extracting the total RNA. Total RNA extracted from non-transgenic Taipei 309 seedlings of the same age was used as a control. Five micrograms of total RNA from each line were separated in a 1% MOPS-formaldehyde agarose gel [20]. After electrophoresis, RNA was blotted to the Hybond-N+[®] membrane (Amersham Biosciences, Piscataway, NJ, USA). The blot was probed with a 759 bp PCR-generated DNA fragment from the *GUS* coding sequence with primers GUSF and GUSR (Supplementary Table 1). DNA probes were labeled with [³²P]-dCTP using Prime-It II[®] random priming kit (Stratagene, Cedar Creek, TX, USA) according to the manufacturer's protocol. Hybridization and washing of the membrane were performed using procedures provided by Amersham. The radioactivity was then detected by a Storm[™] 840 phosphor-imager and analyzed with the Image Quant TL version 2005 software (Amersham). The readings were normalized by the intensity of the ribosomal RNA bands stained with EdBr.

2.5. Verification of intron splicing by RT-PCR

Reverse transcription PCR (RT-PCR) was performed using total RNA from transgenic rice plant lines containing constructs derived from either pRESQ4, pRESQ48 or pRESQ70. Genomic DNA from a pRESQ70 transgenic plant was used to perform PCR as a control to show the position of the non-spliced product. Total RNA from the non-transgenic Taipei 309 rice plant was used as a negative control. RT-PCR was performed with SpliceF and SpliceR primers (Supplementary Table 1) located in the *rubi3* exon 1 region and the *GUS* coding sequence, respectively, using an RT-PCR kit (Applied Biosystems, Foster City, CA) and standard conditions. The amplified PCR products were sequenced with the SpliceR primer. The resulting sequence data were compared with the *rubi3* intron and flanking sequences to verify splicing accuracy in these transgenic lines.

2.6. Base-pairing comparison of leader sequences

A computer program "free2bind" written based on an algorithm described by Starmer et al. [21] was used for analyzing the complementarity between the 5' UTR sequences of the *GUS* transgene mRNAs of various constructs and the rice 17S rRNA [22]. Minimum ΔG° values were calculated using the free-align and free-scan programs [21]. These programs combine experimentally derived thermodynamic parameters with a dynamic programming algorithm that calculates the minimal free energy for hybridizing two RNA sequences at a specific temperature. All the sequence-pairing analyses were performed with the temperature setting at 25 °C. The program is accessible on-line at the following URL: <http://www.unc.edu/~starmer/free2bind/>.

3. Results

3.1. Effects of subsequences from *rubi3*'s second exon on *GUS* expression

Construct pRESQ4 has a 20 nt non-coding leader sequence immediately upstream from the coding sequence, which was carried over from plasmid pAHC25 [6] (Fig. 1). Fourteen nt in this

20 nt sequence were from the original *E. coli uidA* gene 5' UTR [23], and the remaining is a SmaI restriction sequence [6]. In construct pRESQ48, we replaced the 20 nt leader sequence from pRESQ4 with the sequence, ATCCAGATA, which is a mutated version of the first 9 nt of the *rubi3* coding sequence (from exon 2), where the third base was changed from 'G' to 'C' [3]. We observed an increase in transient *GUS* expression from pRESQ48 by about 4-fold over pRESQ4 [3]. The plasmid pRESQ2b was made by placing the above 9 nt sequence in front of the original 20 nt leader sequence in pRESQ4 (Fig. 1). In transient assays, *GUS* expression levels from pRESQ2b were comparable to that from pRESQ4 (Fig. 1). Thus, the enhancement associated with the 9 nt sequence (as seen in pRESQ48) was lost when the original 20 nt leader sequence was between it and the start codon.

3.2. Does the proximity of the translation start codon to the 5' UTR intron affect *GUS* expression?

In the above-described experiment, we noticed that the distances between the 5' UTR intron and the *GUS* translation initiation codon in pRESQ4 and pRESQ2b (20 and 29 nt, respectively) was longer than the one in pRESQ48 (9 nt). We wondered whether the lower *GUS* activities from pRESQ4 and pRESQ2b were due to a longer distance between the intron and the translation initiation codon. To examine this possibility, we created construct pRESQ85, based on pRESQ36 that we reported previously [3]. Plasmid pRESQ36 is a translational fusion of the first 27 nt from the *rubi3* exon 2 coding sequence with the *GUS* coding sequence. The *GUS* activity from construct pRESQ36 was comparable to pRESQ48 in transient assays [3]. We mutated the ATG at the start of the 27 nt *rubi3* coding sequence to ATC in pRESQ36 to create pRESQ85 (Fig. 2). In transient expression assays, the *GUS* activity from pRESQ85 was similar to pRESQ36, and about 4-fold greater than pRESQ4, indicating that the lower expression from pRESQ4 and pRESQ2b was not caused by the greater distance between the translation initiation codon and the 5' UTR intron splicing site (Fig. 2).

3.3. Is the 9 nt from *rubi3* exon 2 necessary for the enhanced gene expression?

In our previous report, we showed that the mutated first 9 nt from *rubi3* exon 2, as in pRESQ48, were sufficient to significantly enhance *GUS* expression levels over pRESQ4 [3]. To study how many of these 9 nt were necessary for the observed expression enhancement, we created pRESQ68 and pRESQ70 constructs from pRESQ48. Construct pRESQ68 had only the first 3 nt (ATC) of the original 9 nt sequence, while all 9 nt were removed to create pRESQ70, eliminating all untranslated sequence between the *rubi3* 5' UTR intron and the *GUS* start codon (Fig. 3, GenBank accession no. EU503043). Transient *GUS* assay results from both constructs were very similar to that of pRESQ48 (Fig. 3), suggesting that none of the *rubi3* coding sequence is necessary for the observed enhancement.

3.4. Analysis of *GUS* expression in transgenic rice plants

To further study *GUS* expression from these constructs at the whole plant level, 24 transgenic plant lines were generated from *Agrobacterium* EHA105 (pTOK47, pCAM70), which contains the *rubi3/GUS* reporter gene construct with no UTR between the 5' UTR intron and the *GUS* start codon. Two transgenic rice plant lines were also recovered from *Agrobacterium* strain EHA105 (pTOK47, pCAM48), which contains the *rubi3/GUS* construct that contains the sequence, ATCCAGATA, between the 5' UTR intron

and the *GUS* start codon. *GUS* transgene copy number was determined by RealTime PCR and confirmed by Southern analysis (data not shown). To study *GUS* expression at both RNA and *GUS* enzyme activity levels, seven T1 plants from line pCAM70E1 were compared with T1 plants from the pCAM48E1 plants. Additionally, we examined two existing plant lines: one (pCAM36E1) derived from pRESQ36, which contained the first 27 nt from *rubis3*'s exon 2 between the *rubis3* 5' UTR intron and the start codon for *GUS* (pCAM36 is referred as pJLU6 in [7]), and a second (pCAM4E1) derived from pRESQ4, which contained the original 20-nt non-coding *GUS* leader sequence (pCAM4 is referred as pJLU2 in [7]). The representative plant lines for each of these constructs were chosen from a pool of single copy plant lines that showed similar *GUS* mRNA levels in an earlier analysis [7]. The results are presented in Fig. 4A and B. Quantification of *GUS* mRNA levels by phosphorimaging revealed very similar expression levels among these lines (Fig. 4A). MUG assays showed that the *GUS* activities in plants from constructs based on pRESQ70, pRESQ48, and pRESQ36 were about 3-fold greater than that of pRESQ4-derived plants (Fig. 4B). The result is consistent with our earlier analysis using transient assays and analysis of transgenic rice plants using constructs derived from pRESQ4 and pRESQ36 [7]. RT-PCR and subsequent sequence analysis of the products indicate that the *rubis3* 5' UTR intron was correctly and well spliced in transgenic rice plants containing various gene constructs (Fig. 5) and that the observed difference in *GUS* activity was not caused by inefficient splicing. The results further confirm that the 3–4-fold enhancement of *GUS* activity from various constructs with exon 2 sequence modifications (pRESQ36, 48 and 70) is translational, as reported earlier [7], and indicate that the relatively low *GUS* expression in pRESQ4 was caused by its non-coding leader sequence.

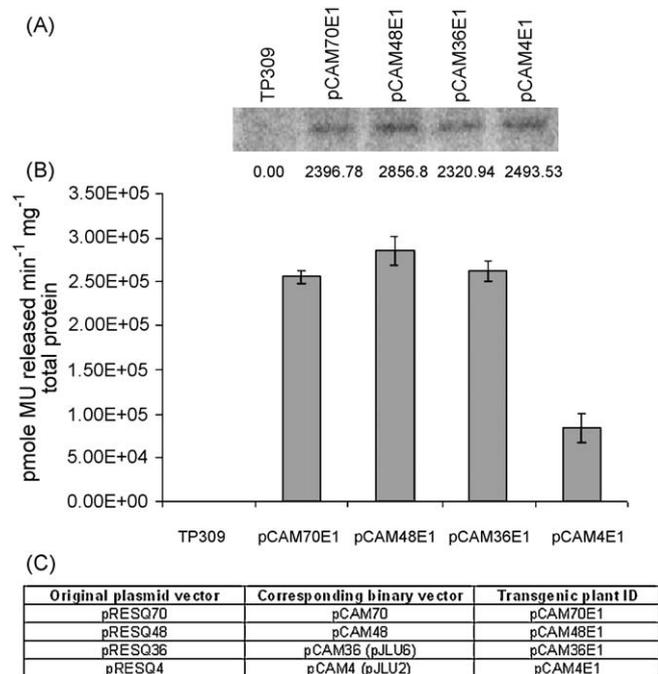


Fig. 4. *GUS* RNA and protein expression in transgenic plants with various gene constructs. With a non-transgenic Taipei 309 plant used as a control (TP309), (A) shows a Northern blot of total RNA, from the same set of transgenic rice plants used in the MUG assays, hybridized with a *GUS* gene probe. The number below each lane represents *GUS* mRNA amounts as determined by phosphorimaging and normalized to quantities of the rRNA. (B) *GUS* enzyme activity in the different transgenic plants. (C) A list of the original plasmid vectors, their corresponding binary vectors used in rice transformation, and the relevant transgenic lines analyzed in the experiment.

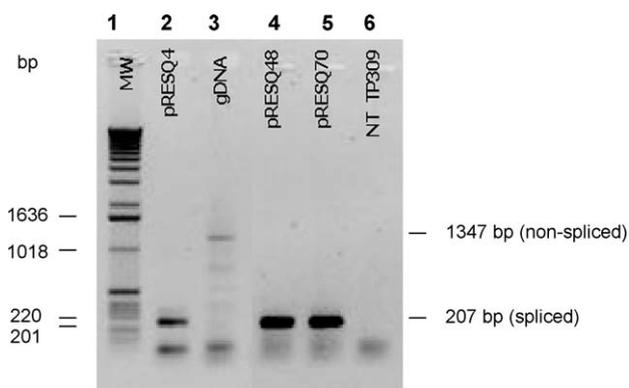


Fig. 5. Verification of intron splicing in transgenic rice plants of various *rubis3* promoter constructs. RT-PCR products of total RNA from plants transformed with pRESQ4, pRESQ48 and pRESQ70, and from a non-transgenic Taipei 309 plant (NT TP309), were fractionated by electrophoresis on a 1% agarose gel as shown in lanes 2, 4, 5 and 6. PCR with genomic DNA was designed to indicate the size of the product from non-spliced *GUS* mRNA (lane 3). Molecular weight standards were loaded in lane 1 and some are marked on the left. The expected sizes of the PCR products, with the intron spliced or unspliced, are indicated on the right. Bands below 100 bp in each lane are primers used in the PCR reactions.

3.5. Removal of the leader sequence from the maize *Ubi1* promoter construct also increased *GUS* expression

To further examine the hypothesis that the non-coding leader sequence in pRESQ4 inhibited translation, we created construct pRESQ95, which is derived from pRESQ29, a maize *Ubi1* promoter-*GUS* construct we had used in our previous studies [3]. The *GUS* gene in pRESQ29 contains a 37 nt non-coding leader sequence [6], of which a segment of 20 nt was identical to that in pRESQ4. The construct pRESQ95 was made identical to pRESQ29 except that the 37 nt leader sequence was completely removed in order to have the first ATG of the *GUS* coding sequence fused right behind the maize *Ubi1* 5' UTR intron, just as in construct pRESQ70 (Fig. 6). Similar to what we observed in the *rubis3* promoter constructs, when the leader sequence was removed from the maize *Ubi1* promoter construct, *GUS* enzyme activity was increased by about 4-fold (Fig. 6). This result further confirms our observation with constructs pRESQ4 and pRESQ70, and suggests that the *GUS* leader sequence have an inhibitory role on gene expression at the translational level.

3.6. Complementarity between non-coding leader sequences and the rice 17S ribosomal RNA negatively correlates with the translational enhancement

From the above experiments, and the results reported previously [3], we conclude that the *rubis3* promoter together with its 5' UTR intron (and likely other plant polyubiquitin promoters such as the maize *Ubi1* with their 5' UTR introns) has a potential to achieve high levels of transgene expression, but the expression could be affected by non-coding leader sequences between the intron and the start codon. As indicated in this report, the *GUS* leader sequences used in the study had a substantial inhibitory effect on gene expression at the translational level. Since all of our experiments were performed in rice, to further gain insight into the observed translational inhibition, we compared the rice 17S ribosomal RNA sequence (GenBank accession no. X00755) with the non-coding leader sequences used in our constructs, and calculated the minimum ΔG° values for best pairing as determined by the "free2bind" software [21]. Although each construct has certain pairing, the constructs pRESQ4 and pRESQ29 showed the strongest pairing, or lowest ΔG° values, between the non-coding

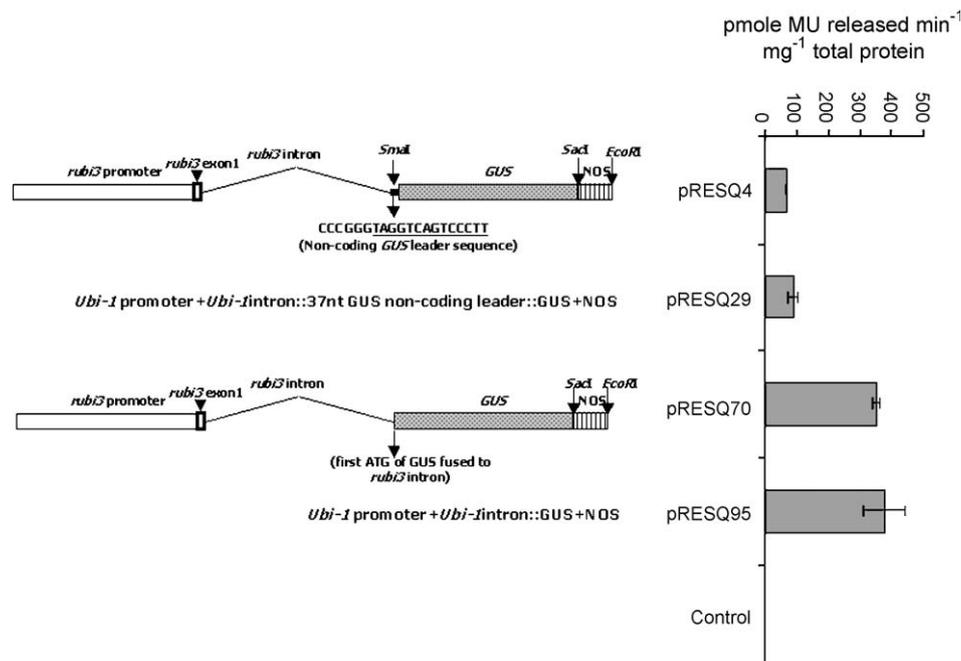


Fig. 6. Removal of the *GUS* leader sequence in a maize *Ubi1* promoter construct enhanced *GUS* gene expression. The *GUS* non-coding leader sequences in plasmid constructs pRESQ4 and pRESQ29 were removed and fusion constructs pRESQ70 and pRESQ95 were made in a way that the *GUS* coding sequence was directly fused to the intron 3' splice site. Each of these constructs was co-bombarded with pJD313 plasmids to the rice suspension cells. The graph shows the *GUS* enzyme activity of each construct normalized to LUC activity. The control cells were bombarded with pJD313 only.

Table 1

Correlations among potential base-pairing between the leader sequences and the rice 17S rRNA and transient expression of *GUS* gene constructs.

| Plasmid construct | Exon 2 untranslated sequence | Lowest minimum ΔG° kcal/mol | No. of 17S rRNA sites with $\Delta G^\circ < -13^a$ | Fold change in expression ^b |
|-------------------|--|--|---|--|
| pRESQ29 | GTCTGACTCTAGAGGATCCCCGGGTAGGTCAGTCCCTT | -18.98 | 14 | -4 |
| pRESQ4 | CCCGGGTAGGTCAGTCCCTT | -18.89 | 8 | -4 |
| PREQ101 | CCCGGG | -13.17 | 1 | 0 |
| pRESQ85 | ATCCAGATATTCGTTAAGACCCTCACT | -12.32 | 0 | 0 |
| pRESQ48 | ATCCAGATA | -10.71 | 0 | 0 |

^a "No. of 17S rRNA sites with $\Delta G^\circ < -13$ " column shows that both pRESQ29 and pRESQ4 have potential to bind to rice 17S rRNA at several locations, in contrast to pRESQ48, pRESQ85, and pRESQ101. "Minimum ΔG° " is an indication of how strong the optimal binding between the two RNA molecules is. The lower the value is, the stronger the binding between the two molecules.

^b Fold change of pRESQ29 is in comparison to pRESQ95 whereas change of pRESQ48, 85, or 101 is relative to pRESQ70. Neither pRESQ95 nor pRESQ70 has exon 2 untranslated sequence.

leader sequences and the 17S rRNA (Table 1). Most strikingly, there are multiple sites (8 for pRESQ4 or 14 for pRESQ29) in rice 17S rRNA that could potentially bind to the leader sequences of these constructs with minimum ΔG° values lower than -13 kcal/mol (Supplementary Table 2), while no such strong binding sites were found in constructs pRESQ85 and pRESQ48, which have higher *GUS* expression.

Analysis of the 20 nt leader sequence present in pRESQ4 using the "free2bind" software revealed that all 20 nt are involved in a fair amount of complementarity to certain rice 17S rRNA sequence fragments with ΔG° values lower than -13 kcal/mol (Supplementary Table 2). These 20 nt are also included in the 37 nt leader sequence in pRESQ29, and are attributed to most of the strong binding with rice 17S rRNA (Supplementary Table 2).

To further test the hypothesis that binding between the leader sequences and the 17S rRNA inhibited translation, and to facilitate cloning of "useful" genes for maximal expression under control of *rubi3* promoter, vector pRESQ101 (GenBank accession no. EU503042) was constructed based on the knowledge gained from "free2bind" software analysis. In pRESQ101, only the *Sma*I restriction site, CCCGGG, from the pRESQ4 leader sequence was

retained and the remaining 14 nt were removed (Fig. 7). By doing so, pRESQ101 eliminated 7 out of 8 potential strong binding sites with rice 17S rRNA from pRESQ4 (Table 1). In transient *GUS* assays, the reporter gene expression of pRESQ101 was comparable to that of pRESQ48 and pRESQ70, and about 4-fold higher than pRESQ4 (Fig. 7).

4. Discussion

The *rubi3* 5' regulatory sequence has comparably strong activities in both rice and maize cells [26], and its expression is constitutive [27], suggesting its potential use in plant biotechnology. Our laboratory cloned and reported the *rubi3* 5' regulatory sequence, including its promoter and the 5' UTR exon 1 and intron, and used it as a model system to study IME in monocot plants [3,7,25,27]. When we included the 5' UTR intron, we observed approximately 2-fold enhancement of transcription, but over 20-fold enhancement of mRNA accumulation in rice suspension cells, indicating that post-transcriptional events play the most important role in IME [24,25]. In our earlier report [3], we demonstrated that gene expression is even further enhanced when the first 9 nt of

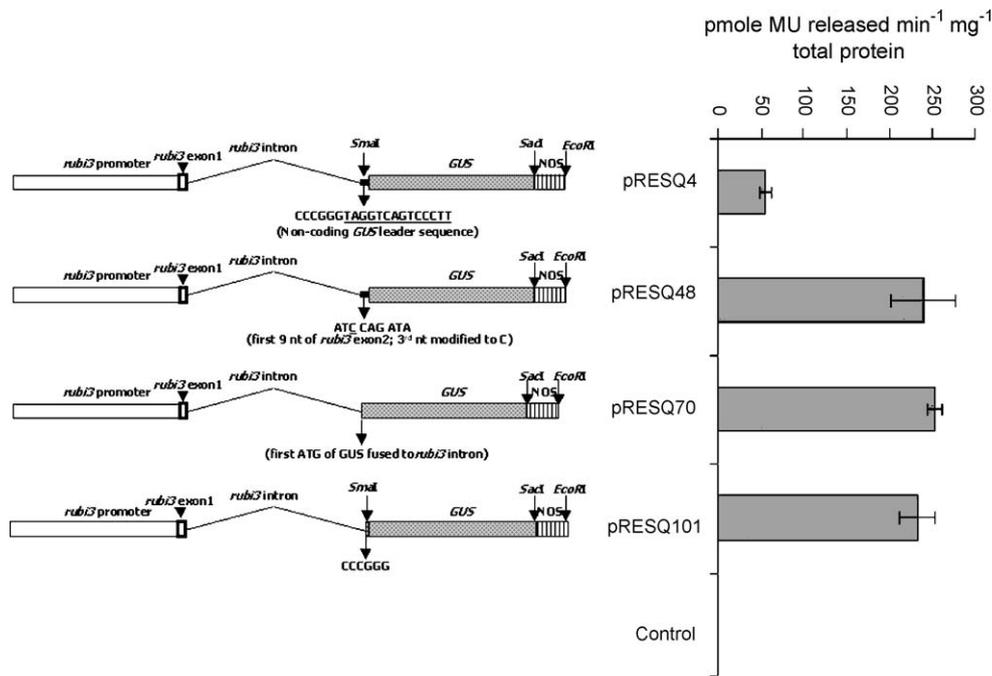


Fig. 7. GUS activity of construct pRESQ101 in comparison with other constructs. The pRESQ101 has 6 nt (*Sma*I restriction site) between *rubi3* intron and the ATG initiation codon of the *GUS* coding sequence. Each of these constructs was co-delivered with pJD313 plasmids to rice suspension cells. The graph shows GUS enzyme activity of each construct normalized to LUC activity. Control cells were bombarded with pJD313 only.

coding sequence (original or mutated) from *rubi3* exon 2 were fused to the *GUS* coding sequence (constructs pRESQ42, 46, 47 and 48). In this report, we showed that the observed enhancement is more likely to be attributable to the removal of the leader sequence carried over from the original *GUS* gene construct, and points to adverse affects leader sequences can have on translation (Fig. 1).

It has been reported that potential base pair formation between the small ribosomal subunit (depending on the species, this is either a 17S or an 18S rRNA and the labels are often used interchangeably) and 5' UTR mRNA leader sequences could affect translation efficiency both positively and negatively [9,28–31]. In one such report in plants, Akbergenov et al. observed *in vitro* and *in vivo* enhancement of translational efficiency of the *GUS* mRNA when sequences complementary to the rice 17S rRNA were introduced into the constructs' 5' untranslated region [30]. Specifically, the enhancement was attributed to base-pairing between the leader sequences and rice 17S rRNA nucleotides 1105–1114 and 1115–1124. While these results may appear to contradict what we are currently reporting, they do not. The strong base-pairing between the 17S rRNA and the leader sequences (present in pRESQ4 and pRESQ29 constructs) that we are proposing to inhibit translation does not include either of these two regions (Supplementary Table 2). Furthermore, one region in rice 17S rRNA that has extensive complementarity with the leader sequences in pRESQ4 and pRESQ29 is homologous to a region in the mouse 18S rRNA that has been shown to bind to mRNA [29]. Our data indicate an inverse correlation between complementarity of the leader sequences with specific sequences within the rice 17S rRNA and *GUS* gene expression, and are consistent with observations in mouse [28,29]. It has been proposed that ribosomal subunits play a regulatory role by mediating interactions between mRNAs and components of translational machinery to determine the translation initiation efficiency of individual mRNAs [9].

It is noticeable that in both pRESQ70 and pRESQ95 the ATG initiation codon in the *GUS* gene is immediately behind the 3' splicing site of the 5' UTR intron. This is identical to the structural organization of the native rice *rubi3* and maize *ubi1* polyubiquitin

genes, and both appear to be in the optimal conformation to lead to their maximal expression. The highest translational enhancement could be related to intron splicing [3,25]. Nott et al. [32] observed that spliced mRNA produced greater quantities of protein per mRNA molecule than the non-spliced, but otherwise identical, mRNA in mammalian cells. The phenomenon is correlated to the enhanced polysome association with the mRNA in cytoplasm, and can be attributed to shuttling components of exon junction complex, which accompany the spliced mRNA from nucleus to cytoplasm. Sanford et al. [33] observed a similar effect and identified the shuttling splicing factors that were associated with actively translating ribosomes and stimulated translation. Even with gene constructs that contain the *GUS* leader sequence (pRESQ4 and its intronless counterpart, pPSRG30), we observed moderate translational enhancement mediated by intron splicing in callus and root cells, although the enhancement appeared to be tissue dependent [7,25]. We have demonstrated that the *rubi3* polyubiquitin promoter (with its 5' UTR intron) enhances gene expression at transcriptional, RNA processing, and translational levels [7,25]. These results provide insights regarding why polyubiquitin gene promoters are among the strongest promoters in plants.

In transgene construction, it is common for 5' UTR sequences to remain with the original gene during the ligation process. Our current results indicate that these extra sequences could reduce translation efficiency of the transgene. With modern computational tools such as "free2bind", one could easily predict the binding of such 5' UTR sequences to that of the small rRNA subunit and avoid costly mistakes.

Acknowledgements

Plasmid from Drs. P.H. Quail (pAHC25), J. Sheen (pJD313), and R. Wu (*Act1* gene) and T1 seeds of transgenic rice plants from Dr. J. Lu are gratefully acknowledged. We thank T.D. Tuong and K. Siravuri for technical assistance. We are thankful to Drs. M. Vouk and L. Ponnala for initial secondary RNA complementation predictions

and helpful discussions. We also thank Dr. R. Dewey for critical reading of this manuscript and suggestions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2009.08.006.

References

- [1] J. Callis, R.D. Vierstra, Ubiquitin and ubiquitin genes in higher plants, *Oxford Survival Plant Molecular Cell Biology* 6 (1989) 1–30.
- [2] M. Schelesinger, U. Bond, Ubiquitin genes, *Oxford Survival Eukaryotic Genes* 4 (1987) 77–89.
- [3] E. Sivamani, R. Qu, Expression enhancement of a rice polyubiquitin gene promoter, *Plant Molecular Biology* 60 (2) (2006) 225–239.
- [4] M.J. Cornejo, D. Luth, K.M. Blankenship, O.D. Anderson, A.E. Blechl, Activity of a maize ubiquitin promoter in transgenic rice, *Plant Molecular Biology* 23 (3) (1993) 567–581.
- [5] S.J. Streatfield, M.E. Magallanes-Lundback, K.K. Beifuss, C.A. Brooks, R.L. Harkey, R.T. Love, J. Bray, J.A. Howard, J.M. Jilka, E.E. Hood, Analysis of the maize polyubiquitin-1 promoter heat shock elements and generation of promoter variants with modified expression characteristics, *Transgenic Research* 13 (4) (2004) 299–312.
- [6] A.H. Christensen, P.H. Quail, Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants, *Transgenic Research* 5 (3) (1996) 213–218.
- [7] J. Lu, E. Sivamani, K. Azhakanandam, P. Samadder, X. Li, R. Qu, Gene expression enhancement mediated by the 5' UTR intron of the rice *rub13* gene varied remarkably among tissues in transgenic rice plants, *Molecular Genetics and Genomics* 279 (2008) 563–572.
- [8] D.R. Gallie, Posttranscriptional regulation of gene-expression in plants, *Annual Review of Plant Physiology and Plant Molecular Biology* 44 (1993) 77–105.
- [9] V.P. Mauro, G.M. Edelman, The ribosome filter hypothesis, *Proceedings of the National Academy of Sciences of the United States of America* 99 (19) (2002) 12031–12036.
- [10] D.R. Gallie, The 5'-leader of tobacco mosaic virus promotes translation through enhanced recruitment of eIF4F, *Nucleic Acids Research* 30 (15) (2002) 3401–3411.
- [11] T. Kato, Y. Shirano, T. Kawazu, Y. Tada, E. Itoh, D. Shibata, A modified β -glucuronidase gene: sensitive detection of plant promoter activities in suspension-cultured cells of tobacco and rice, *Plant Molecular Biology Reporter* 9 (4) (1991) 333–339.
- [12] C.A. Helliwell, J.C. Gray, The sequence surrounding the translation initiation codon of the pea plastocyanin gene increases translational efficiency of a reporter gene, *Plant Molecular Biology* 29 (3) (1995) 621–626.
- [13] E.R. LaVallie, J.M. McCoy, Gene fusion expression systems in *Escherichia coli*, *Current Opinions in Biotechnology* 6 (5) (1995) 501–506.
- [14] D. Hondred, J.M. Walker, D.E. Mathews, R.D. Vierstra, Use of ubiquitin fusions to augment protein expression in transgenic plants, *Plant Physiology* 119 (2) (1999) 713–724.
- [15] J. Yon, M. Fried, Precise gene fusion by PCR, *Nucleic Acids Research* 17 (12) (1989) 4895–14895.
- [16] G. An, P.R. Ebert, A. Mitra, S.B. Ha, in: S.B. Gelvin, R.A. Schilperoort, D.P.S. Verma (Eds.), *Plant Molecular Biology Manual*, vol. A3, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988, pp. 1–19.
- [17] E.E. Hood, S.B. Gelvin, L.S. Melchers, A. Hoekema, New agrobacterium helper plasmids for gene-transfer to plants, *Transgenic Research* 2 (4) (1993) 208–218.
- [18] S.G. Jin, T. Komari, M.P. Gordon, E.W. Nester, Genes responsible for the super-virulence phenotype of *Agrobacterium tumefaciens* A281, *Journal of Bacteriology* 169 (10) (1987) 4417–4425.
- [19] CAMBIA, Transformation Protocol for Rice—Abbreviated, as posted on Bioforge.net (2005).
- [20] E. Sivamani, H. Huet, P. Shen, C.A. Ong, A. de Kochko, C. Fauquet, R.N. Beachy, Rice plant (*Oryza sativa* L.) containing Rice tungro spherical virus (RTSV) coat protein transgenes are resistant to virus infection, *Molecular Breeding* 5 (2) (1999) 177–185.
- [21] J. Starmer, A. Stomp, M. Vouk, D. Bitzer, Predicting shine-dalgarno sequence locations exposes genome annotation errors, *PLoS Computational Biology* 2 (5) (2006) e57.
- [22] F. Takaiwa, K. Oono, M. Sugiura, The complete nucleotide-sequence of a rice 17S ribosomal-RNA gene, *Nucleic Acids Research* 12 (13) (1984) 5441–5448.
- [23] R.A. Jefferson, T.A. Kavanagh, M.W. Bevan, Gus fusions—beta-glucuronidase as a sensitive and versatile gene fusion marker in higher-plants, *EMBO Journal* 6 (13) (1987) 3901–3907.
- [24] P. Samadder, E. Sivamani, J. Lu, X. Li, R. Qu, The 16th Penn State Plant Physiology Symposium—RNA Biology: Novel Insights from Plant Systems, RNA Biology, University Park, PA, 18–20 May 2006, 2006.
- [25] P. Samadder, E. Sivamani, J. Lu, X. Li, R. Qu, Transcriptional and post-transcriptional enhancement of gene expression by the 5' UTR intron of rice *rub13* gene in transgenic rice cells, *Molecular Genetics and Genomics* 279 (2008) 429–439.
- [26] E. Sivamani, R.K. DeLong, R. Qu, Protamine-mediated DNA coating remarkably improves bombardment transformation efficiency in plant cells, *Plant Cell Reports* 28 (2009) 213–221.
- [27] J. Lu, E. Sivamani, X. Li, R. Qu, Activity of the 5' regulatory regions of the rice polyubiquitin *rub13* gene in transgenic rice plants as analyzed by both GUS and GFP reporter genes, *Plant Cell Reports* 27 (2008) 1587–1600.
- [28] M.C.Y. Hu, P. Tranque, G.M. Edelman, V.P. Mauro, rRNA-complementarity in the 5' untranslated region of mRNA specifying the Gtx homeodomain protein: Evidence that base-pairing to 18S rRNA affects translational efficiency, *Proceedings of the National Academy of Sciences of the United States of America* 96 (4) (1999) 1339–1344.
- [29] S.B. Verrier, O. Jean-Jean, Complementarity between the mRNA 5' untranslated region and 18S ribosomal RNA can inhibit translation, *RNA—a Publication of the RNA Society* 6 (4) (2000) 584–597.
- [30] R.Z. Akbergenov, S.S. Zhanybekova, R.V. Kryldakov, A. Zhigailov, N.S. Polimbetova, T. Hohn, B.K. Isakov, ARC-1, a sequence element complementary to an internal 18S rRNA segment, enhances translation efficiency in plants when present in the leader or intercistronic region of mRNAs, *Nucleic Acids Research* 32 (1) (2004) 239–247.
- [31] S.A. Chappell, G.M. Edelman, V.P. Mauro, A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity, *Proceedings of the National Academy of Sciences of the United States of America* 97 (4) (2000) 1536–1541.
- [32] A. Nott, H. Le Hir, M.J. Moore, Splicing enhances translation in mammalian cells: an additional function of the exon junction complex, *Genes & Development* 18 (2) (2004) 210–222.
- [33] J.R. Sanford, N.K. Gray, K. Beckmann, J.F. Caceres, A novel role for shuttling SR proteins in mRNA translation, *Genes & Development* 18 (7) (2004) 755–768.