ORIGINAL PAPER

Transcriptional and post-transcriptional enhancement of gene expression by the 5′ UTR intron of rice *rubi3* gene in transgenic rice cells

Partha Samadder · Elumalai Sivamani · Jianli Lu · Xianggan Li · Rongda Qu

Abstract Introns play a very important role in regulating gene expression in eukaryotes. In plants, many introns enhance gene expression, and the effect of intron-mediated enhancement (IME) of gene expression is reportedly often more profound in monocots than in dicots. To further gain insight of IME in monocot plants, we quantitatively dissected the effect of the 5′ UTR intron of the rice *rubi3* gene at various gene expression levels in stably transformed suspension cell lines. The intron enhanced the *GUS* reporter gene activity in these lines by about 29-fold. Nuclear run-on experiments demonstrated a nearly twofold enhancement by the 5′ UTR intron at the transcriptional level. RNA analysis by RealTime quantitative RT-PCR assays indicated the intron enhanced the steady state RNA level of the *GUS* reporter gene by nearly 20-fold, implying a strong role of the intron in RNA processing and/or export. The results also implicated a moderate role of the intron in enhancement at the translational level ($\sim$45%). Moreover, results from a transient assay experiment using a shortened exon 1 sequence revealed an important role of exon 1 of *rubi3* in gene expression. It may also hint a divergence in IME mechanisms between plant and animal cells. These results demonstrated transcriptional enhancement by a plant intron, but suggested that post-transcriptional event(s) be the major source of IME.

Keywords EJC · Exon · Gene expression · IME · Nuclear run-on

P. Samadder · E. Sivamani · J. Lu · R. Qu (✉)
Department of Crop Science,
North Carolina State University, 1126A Williams Hall,
P.O. Box 7620, Raleigh, NC 27695, USA
e-mail: rongda_qu@ncsu.edu

X. Li
Syngenta Biotechnology, Inc., Research Triangle Park,
Raleigh, NC 27709, USA

Present Address:
P. Samadder
Department of Immunobiology, College of Medicine,
University of Arizona, Tucson, AZ 85724, USA

Present Address:
J. Lu
Department of Plant Biology,
North Carolina State University, 4210 Gardner Hall,
P.O. Box 7612, Raleigh, NC 27695, USA

Introduction

Most genes in eukaryotes are interrupted by at least one non-coding intervening sequence (intron), which is present in nascent pre-mRNA and absent in mature RNA. The role of introns to enhance gene expression has been well documented in various organisms (Buchman and Berg 1988; Okkema et al. 1993; Callis et al. 1987; Luehrsen and Walbot 1991; Rose and Last 1997). It has been suggested that transcription, intron splicing, and RNA processing (capping and 3′ processing) are integrated process and the reactions affect one another (Proudfoot et al. 2002). Studies in animal cells revealed that the enhancement is achieved through a combinatorial mechanism that increases transcription, mRNA accumulation, and translation (Furger et al. 2002; Nott et al. 2003, 2004) while not substantially affecting mRNA stability (Nott et al. 2003). Upon intron splicing, a protein complex, named exon junction complex (EJC), was deposited 20–24 nucleotides upstream of the exon junction site (Le Hir et al. 2000). It has been shown

that the EJC binding, rather than the splicing itself, is the primary determinant for the enhanced gene expression mediated by the intron (Wiegand et al. 2003; Nott et al. 2004). Some EJC proteins are shuttle proteins and accompany the spliced mRNA from the nucleus to the cytosol (Kataoka et al. 2000, 2001). Although such an EJC has not been demonstrated in plants, several conserved homologs of the human EJC proteins were identified in *Arabidopsis* genome analysis, and those proteins were localized at or associated with the nucleolus (Pendle et al. 2005).

Although mRNA processing occurs across kingdoms, animal and yeast introns are not usually correctly spliced in plants. Thus, while the process of pre-mRNA splicing may be conserved (Simpson and Filipowicz 1996), a difference in intron sequences, such as AU content (Goodall and Filipowicz 1989; Lambermon et al. 2000), and/or in the 5′ splice site selection mechanism (Egoavil et al. 1997; McCullough and Schuler 1997) probably prevent their correct processing in plant cells. Introns are abundant in plant genomes. It is estimated that about 80% of plant nuclear genes contain introns (Goodall et al. 1991; Initiative 2000; Reddy 2001). Plant introns have a high AU content, which is critical for their efficient splicing (Goodall and Filipowicz 1989; Luehrsen and Walbot 1994). Many introns have been shown to increase gene expression in plants (Callis et al. 1987; Luehrsen and Walbot 1991; Luehrsen et al. 1994; Simpson et al. 1992; Bourdon et al. 2001; Clancy and Hannah 2002; Rose 2002). This intron-mediated enhancement (IME) of gene expression within plants is reportedly often higher in monocots (from dozens to over 100-fold) than in dicots (normally two to tenfold) (Simpson and Filipowicz 1996; Clancy and Hannah 2002; Rose 2002), and is dependent on both the position and the sequence of the intron (Bourdon et al. 2001; Clancy and Hannah 2002; Rose 2002, 2004). The enhancement could even be affected by the coding sequence linked with an intron (Rethmeier et al. 1997, 1998). As in animal cells (Lu and Cullen 2003; Nott et al. 2003), IME in plants is generally associated with increased accumulation of mRNA (Callis et al. 1987; Rose and Last 1997; Rose 2004). It was observed that the corresponding protein expression, as measured by the reporter enzyme activity, was often elevated more than the increased mRNA accumulation, implying an additional enhancing effect at the translation level (Mascarenhas et al. 1990; Bourdon et al. 2001; Rose 2004). In *Arabidopsis*, an observation that mRNA accumulation of the reporter gene declined with distance of the intron from the promoter made the author to speculate a possible link between the IME and transcription (Rose 2004) although direct experimental evidence by the nuclear run-on in a previous report did not support the model (Rose and Beliakoff 2000). By analyzing the *Arabidopsis PAT1* intron 1, it was suggested that IME was independent of unique intron sequences, and splicing was not required (Rose and Beliakoff 2000). In contrast to the observation in the dicot system, research using the maize *Sh1* gene intron 1 indicated that the intron splicing was essential for gene expression enhancement, and a U-rich motif of the intron sequence increased expression without affecting splicing (Clancy and Hannah 2002). Findings were so divergent, and sometimes controversial, that Clancy and Hannah (2002) suggested "fundamental different mechanisms" may exist in plant IME among various introns, especially between dicots and monocots. The IME studies have been confounded by the facts that only a few plant introns have been intensively studied on the enhancement mechanism, and that most of the research approaches in dicots used stably transformed plants, and the work reported in monocots was more often based on transient assays on suspension cells or protoplasts.
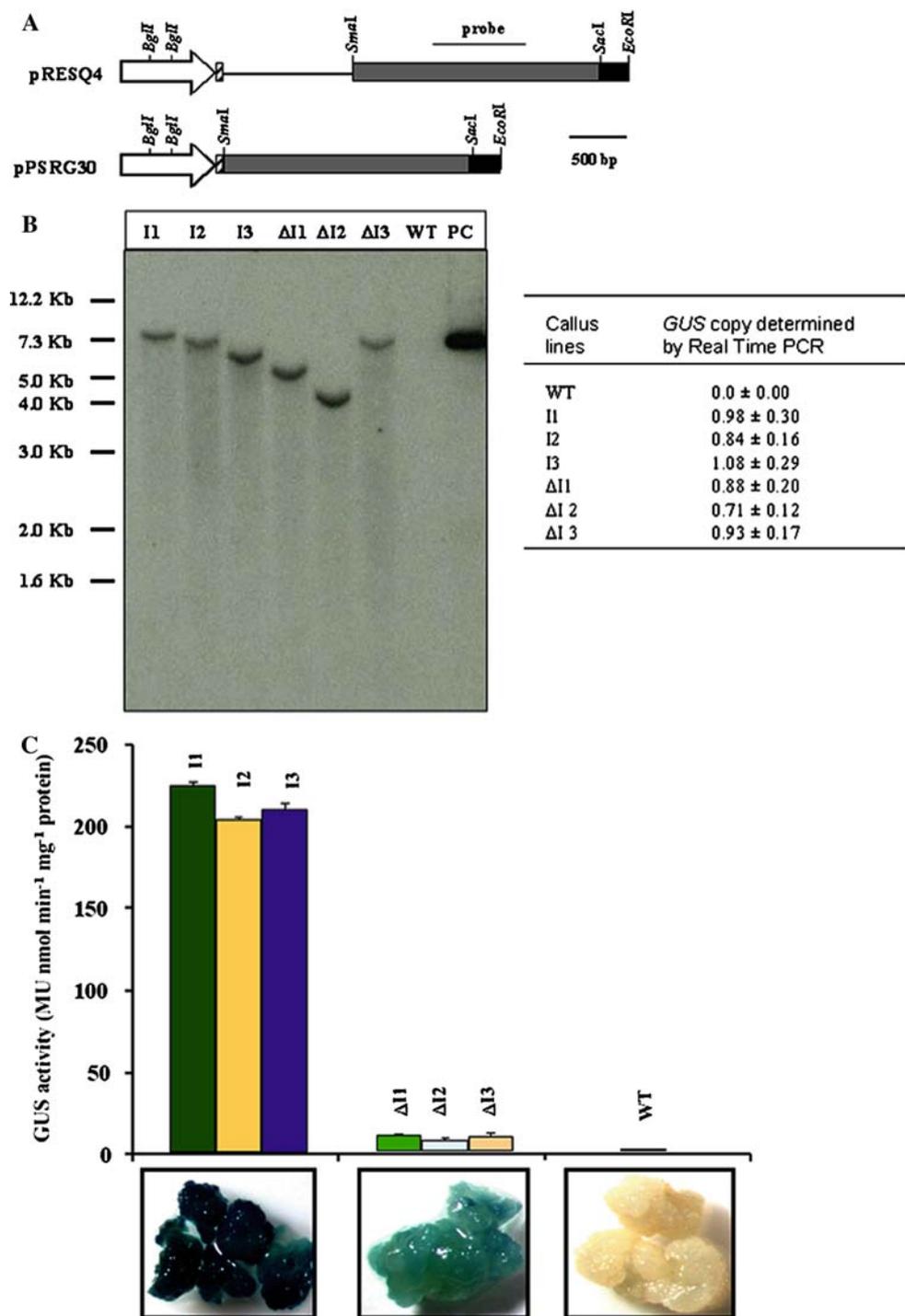
We recently cloned a rice polyubiquitin gene, *rubi3* (Sivamani and Qu 2006), which has a promoter of 0.9 kb, a non-coding 67 bp exon 1, a 5′ UTR intron of 1.1 kb, and a coding region of 1.1 kb (exon 2). The intron displayed a strong IME (∼20-fold) in transient expression assays (Sivamani and Qu 2006). In our query to understand the mechanism of IME in monocots, we used the *rubi3* 5′ regulatory sequence (promoter, exon 1, and 5′ UTR intron), *GUS* reporter gene, and stably transformed rice cell lines as an experimental system to dissect IME at various gene expression levels. In nuclear run-on experiments, we observed a nearly twofold enhancement at transcriptional level. We also employed RealTime quantitative RT-PCR technique to accurately estimate the *GUS* mRNA levels in the transgenic cell lines. Moreover, a transient assay using a gene construct with a shortened exon 1 in our experiments hinted a possible divergence in IME mechanisms between plant and animal cells in the mode of action of a potential plant EJC (exon junction complex) from its counterpart in animal cells.

Materials and methods

Gene construction and plasmid preparation

Construction of plasmid pRESQ4 (Fig. 1a) was reported before (Sivamani and Qu 2006). For making a construct without the 5′ UTR intron and otherwise identical to pRESQ4, a 0.9 kb PCR product of the *rubi3* promoter was cloned in pCR2.1 (Invitrogen, Carlsbad, CA, USA). A *Sma*1 and *Xba*1 fragment of pCR2.1 was sub-cloned in pRESQ8 (Sivamani and Qu 2006) to create pPSRG30 (Fig. 1a). For making a construct with a shortened exon 1 (designated as SE) based on pRESQ4, three-primer PCR (Yon and Fried 1989; Sivamani and Qu 2006) using the

Fig. 1 Intron effect on *GUS* transgene expression in transgenic rice suspension lines. **a** Two *GUS* constructs (pRESQ4 and pPSRG30) used in the experiments: the *rubi3* promoter, its 5′ UTR exon 1 (67 bp), the 5′ UTR intron, the *GUS* coding sequence and the *nos* terminator are presented by *white arrow*, *hatched box*, *thin line*, *gray* and *dark boxes*, respectively. Some restriction sites, including *Bgl*I, which were used for digestion of genomic DNA in Southern analysis, are also shown. The fragment of the *GUS* coding sequence used as a probe for Southern and Northern analyses is also indicated. **b** Trangene copy number determination: the *left panel* is Southern analysis with genomic DNA digested with *Bgl*I. The *right panel* is the RealTime PCR results. **c** GUS enzyme activity in transgenic suspension lines: the mean and standard deviation of GUS enzyme activity in each transgenic suspension line, as a result of MUG assays of three replicates, are presented. Representative GUS histochemical assay results from the two groups of suspension lines and a non-transgenic suspension line (WT) are also shown. I1, 2, 3: transgenic lines from intron-containing construct pRESQ4; ΔI1, 2, 3: transgenic lines from construct pPSRG30 with no intron

Callus lines	*GUS* copy determined by Real Time PCR
WT	0.0 ± 0.00
I1	0.98 ± 0.30
I2	0.84 ± 0.16
I3	1.08 ± 0.29
ΔI1	0.88 ± 0.20
ΔI 2	0.71 ± 0.12
ΔI 3	0.93 ± 0.17

intron-containing pRESQ4 as a template was performed to shorten the exon 1 sequence to 14 bp by fusing the 5′ terminal 7 bp to the 3′ terminal 7 bp of the exon 1 in the construct. The primers used for the PCR were as follows:

SE-1F: 5′-CGCGTGGCGGTGACCGCACCGGATG-3′;
SE-2F: 5′-CCATTAATCGCATAGCGACGTCAAGG TAACTAATCA-3′, and
SE-3R: 5′-CTAACTCGCACGGGATCCTCCTACG-3′.

To make an intronless construct with the same shortened exon 1 (designated as SEΔI) as a control to SE, three-primer PCR using SE-1F, SE-2F and primer Sewoi-3R (5′-CACAAACGGTGATACGTACACTTTTCCC-3′) was carried out to replace exon 1 in pPSRG30 with the same shortened exon 1 (14 bp). The Platinum® *Taq* DNA polymerase High Fidelity (Invitrogen) was used to make both SE constructs. The reactions were performed under the

following conditions: 94°C for 3 min, followed by 30 cycles at 94°C for 2 min, 68°C for 2 min, and 70°C for 2 min, and then at 72°C for 7 min. The sequences of all the DNA fragments obtained by PCR were verified by sequencing analysis. The plasmids used for transformation experiments and nuclear run-on assays were purified using Qiagen® plasmid Midi kit (Qiagen, Valencia, CA, USA).

Development of transgenic rice callus lines

GUS constructs from pRESQ4 and pPSRG30 were inserted into the polylinker region of the binary vector pCAM-BIA1300, and the resultant plasmids were mobilized into *Agrobacterium tumefaciens* strain EHA105 (pTOK47), respectively (Lu 2006). The mature seed-derived rice calli (cv. Taipei 309) were transformed according to the reported procedures (Hiei et al. 1994; Azhakanandam et al. 2000) with modifications (Lu 2006) using hygromycin B (A.G. Scientific, San Diego, CA, USA) selection. For each construct, approximately 20 individual hygromycin B-resistant, GUS-positive callus lines were obtained and their *GUS* transgene copy number determined by RealTime PCR (Ingham et al. 2001) using TaqMan Universal PCR Master Mix (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. Seven lines from pRESQ4 and 10 lines from pPSRG30 were determined as single *GUS* copy lines (with readings between 0.4 and 1.2). Three callus lines from each construct, with *GUS* transgene copy number readings below or slightly above 1 (Fig. 1b), were chosen to develop transgenic suspension lines. They were not chosen based on their expression levels. These lines were I1, I2 and I3 from the intron-containing pRESQ4 construct and ΔI1, ΔI2, and ΔI3 from the intronless pPSRG30. Wild-type Taipei 309 rice suspension culture, provided by Syngenta Biotechnology, was used as a control. All the suspensions were maintained in SZ medium (Zhang et al. 1998) in dark at 25°C and subcultured once a week. The culture medium for transgenic lines contained 50 mg/l hygromycin B.

Isolation of genomic DNA and Southern analysis

Genomic DNA from suspension cells was isolated using a procedure (Dellaporta et al. 1983) with addition of DNase-free RNase A treatment (Sigma, 0.5 mg/ml, 37°C, 30 min). The extracted DNA was digested with restriction enzyme *Bgl*I (Promega, Madison WI, USA) over a period of 18 h and separated on a 0.9% (w/v) agarose gel. DNA was blotted to the Hybond–N+® nylon membrane (Amersham Bioscience, Piscataway, NJ, USA) and probed with a 759 bp PCR-amplified DNA fragment located in the middle of the *GUS* coding sequence (Fig. 1), which was labeled with [32P]-dCTP (Amersham) using Prime-It II® kit (Stratagene,

Cedar Creek, TX, USA). The primers used for *GUS* DNA amplification were Gus1F (5′-CAACGAACTGAACTGG-CAGA-3′) and Gus1R (5′-TTTTTGTCACGCGCTAT-CAG-3′). Hybridization was performed using MiracleHyb® (Stratagene) hybridization solution based on the manufacturer's instructions, and the blot was exposed to a Kodak BioMax® MS film (Kodak Eastman, Rochester, NY, USA) for 6–18 h for autoradiography.

GUS enzyme assays

MUG assays were performed for both stable and transient expression experiments as described (Gallagher 1992; Sivamani and Qu 2006) using a FLUO-star BMG fluorometer (BMG Lab Technologies, Durham, NC, USA). In bombardment transient expression assays, to minimize variations from shot to shot, the same mole amount of plasmid DNA was applied, and a firefly luciferase gene (*LUC*) construct was included as an internal control to normalize the GUS data (Sivamani and Qu 2006). At least three replicates were performed in all the GUS assays and the means and standard deviations were calculated and presented. Histochemical GUS assay was performed using X-Gluc as a substrate according to the reported protocols (Jefferson et al. 1987; Klein et al. 1988).

Isolation of nuclei

Procedure used for isolation of nuclei from plant leaf (Ingelbrecht and de Carvalho 1992; Blockland et al. 1994) was employed, with some modifications, to isolate nuclei from transgenic rice suspension lines. In brief, 15 g of fresh transgenic suspension cells (3 days after subculture) were ground with a small volume of ice-cold, modified Honda buffer (Ingelbrecht and de Carvalho 1992). The ground cells were transferred to a refrigerated blender (Waring, Model 51BL31, Torrington, CT, USA) with 10 ml Honda buffer in it, and blended 3 times, 30 s each. The ground cells were suspended in 10 vol. of Honda buffer and passed twice through a Miracloth® filter (Calbiochem, La Jolla, CA, USA) to remove the cell debris. The filtrate was centrifuged at 2,800 g for 10 min at 4°C, and the resuspended pellet was placed on a discontinuous gradient (40, 60, 80%) of percoll® (MP Biomedicals, Aurora, OH, USA) containing 0. 45 M sucrose and a 2 M sucrose bottom layer, buffered by 250 mM Tris–HCl, pH 8.5, and 10 mM MgCl$_2$ (Ingelbrecht and de Carvalho 1992). The percoll® gradient was centrifuged at 4,000*g* for 30 min to isolate intact nuclei, which were recovered at the interface between the 80% percoll and the bottom layer, and in the pellet. The isolated nuclei were washed with a nuclei wash buffer (NWB), resuspended in a nuclei resuspension buffer (NRB), frozen in liquid N$_2$, and stored at

−70°C until use, as reported by Ingelbrecht and de Carvalho (1992).

Nuclear run-on assays

Nuclear run-on assays were performed according to the published protocol (Blockland et al. 1994) with minor modifications as described below. About 3×10^6 nuclei were used for each nuclear run-on reaction in a volume of 300 μl transcription buffer. The reaction was started by addition of ATP, CTP and GTP, together with 250 μCi of [α-^{32}P]-UTP (3,000 Ci mmol^{-1}, Amersham), and incubated at 28°C for 60 min. Then, tRNA, buffer C (200 mM HEPES, pH 7.6, 5 mM MgCl$_2$, and 5 mM CaCl$_2$) and RQ1 RNase-free DNase (Promega) were added to the reaction mixture and incubated at 37°C for another 30 min to remove the remaining DNA. The reaction was stopped by addition of buffer D (100 mM Tris–HCl, pH 7.5, 50 mM EDTA and 10% SDS) followed by a proteinase K treatment. The mixture was extracted with phenol–chloroform–isoamylalcohol (25:24:1), and the nucleic acids were recovered by ethanol precipitation. Finally the nucleic acids were passed through a Sephadex G-50 quick spin column (Roche, Indianapolis, IN, USA) and the purified, ^{32}P-labeled transcripts were collected. An aliquot was taken for measuring the specific activity using a scintillation counter. The specific activity varied from 10^7 to 4×10^7 c.p.m. per batch of nuclei (3×10^6). The labeled RNAs were used to hybridize a nylon membrane that was previously blotted with 5 μg of λ-phage DNA (negative control), linearized plasmid DNA containing rice *Act1*, or *GUS* coding sequence, respectively, in various slots using a slot blot apparatus (Hoefer Pharmacia, San Francisco, CA, USA). Prehybridization was performed in 4 ml of hybridization buffer (6× SSPE, 5× Denhardt's solution, 0.5% SDS and 0.1 mg/ml salmon sperm DNA) for overnight at 65°C. Each hybridization contained an equal amount of radioactive RNA (10^7 c.p.m.) and the hybridization was performed in the same hybridization buffer containing 10% dextran sulfate (Sigma, St. Louis, MO, USA) for 72 h at 65°C. The membrane was washed with 2× SSC and 0.1% SDS at room temperature, followed by washes with 1× SSC with 0.1% SDS; and 0.1× SSC with 0.1% SDS, once each at 65°C for 30 min, and the signals were quantified with the ImageQuant 5.2 software after scanning with the Storm™ 840 Imaging system (Molecular Dynamics, Sunnyvale, CA, USA). In data analysis, the background counts were subtracted from the counts of *GUS* and *Act1* transcripts, and the *Act1* counts were used to normalize the *GUS* counts from each line. The nuclear run-on experiments were performed three times for each suspension line, and the average value from the last two experiments were analyzed and presented. The significance of difference between the two constructs was deter-

mined by a two-sample unequal variance *t* test using the software package SAS® 9.1 (SAS Institute, Cary, NC, USA).

RNA isolation and northern analysis

Transgenic rice suspension cells were collected on a Whatman #1 filter paper followed by microfuge centrifugation (3,000*g* for 2 min), and stored in liquid nitrogen. Total RNA from each cell line was isolated using RNeasy Plant Midi Kit (Qiagen). Total RNA (20 μg per sample) was size separated on a 1% MOPS–formaldehyde agarose gel by electrophoresis (Sivamani et al. 1999) and transferred onto Hybond–N+® nylon membrane (Amersham). *GUS* gene probe was prepared as described above. Hybridization and washings of filters were performed using procedures provided by Amersham. The radioactivity on the washed filters was detected by Storm™ 840 phosphorimager and analyzed by the Image Quant 5.2 software.

RealTime quantitative RT-PCR

The steady-state level of *GUS* mRNA accumulation in transgenic plants was evaluated by RealTime qRT-PCR. The *GUS* and *NOS* primers (see below) used for the qRT-PCR were designed by IDT PrimerQuest tools (http://scitools.idtdna.com/Primerquest/). The primers for the reference gene (rice *Act1*) were first designed by Kim et al. (2003) and further checked using IDT PrimerQuest tools. Total RNA was extracted from suspension lines using RNeasy® Plant Mini Kit. Prior to cDNA synthesis, 1 μg of RNA was treated with RQ1 RNase-free DNase, and reverse transcription was performed in 10 μl reaction containing 0.5 μg total RNA according to the instruction from the manufacturer (Applied Biosystems). The qRT-PCR was performed in a 25 μl total volume containing 10 μl of 10× diluted cDNA, 50 nM each primer and 12.5 μl Power SYBR® Green PCR Master Mix on an ABI 7900 HT Fluorescent Detection System (Applied Biosystems). The reaction conditions were as follows: 50°C for 2 min, 95°C for 10 min, followed by 45 cycles of 95°C for 15 s, 60°C for 1 min. The experiments were replicated three or four times, using rice *Act1* gene as a reference (Kim et al. 2003), and the means and standard deviations were calculated. Relative *GUS* mRNA level was calculated as the ratio of the PCR product quantity of the *GUS* gene over that of the rice *Act1* gene. Primers used for this experiment were as follows:

qGUS2F: 5′-ATATTGCGCGTTGGCGGTAA-3′,
qNos3R: 5′-CAAGACCGGCAACAGGATTCAA-3′,
qAct6F: 5′-GAAGATCACTGCCTTGCTCC-3′, and
qAct6R: 5′-CGATAACAGCTCCTCTTGGC-3′.

Results

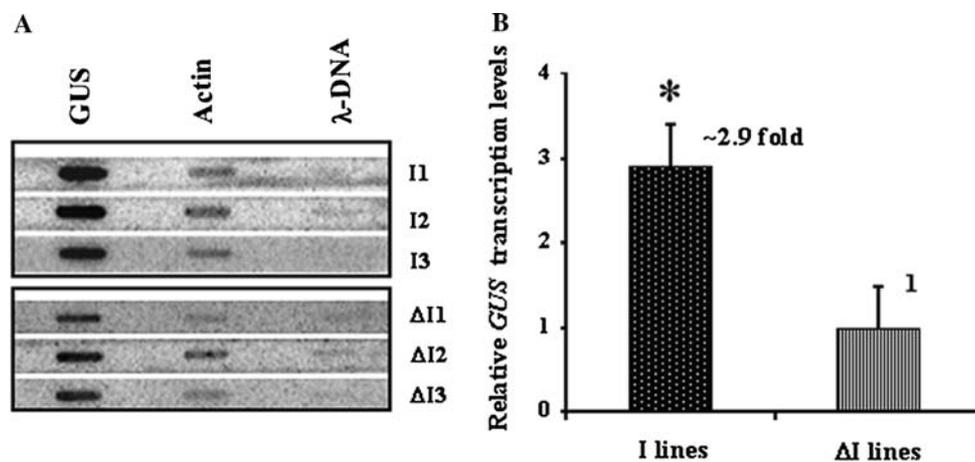### Establishment of transgenic rice suspension lines

The vector pRESQ4 was made to include the *rubi3* promoter, exon 1, intron, *GUS* reporter gene, and *NOS* terminator. An intronless, and otherwise identical, construct, pPSRG30, was also made for the experiments (Sivamani and Qu 2006; Fig. 1a). The two constructs were used to transform rice callus using *Agrobacterium*-mediated approach. To minimize the complicated effects caused by multiple transgene copies, only lines with a single copy of the *GUS* reporter transgene, as determined by a RealTime PCR assay (Ingham et al. 2001) and further supported by Southern analysis (Fig. 1b), were used in the experiments. Three independently transformed callus lines, with a single copy of the *GUS* transgene, from each construct were chosen to develop suspension lines for further study. They were designated as I1, I2, and I3 derived from pRESQ4 (with the intron), and ΔI1, ΔI2, and ΔI3 from pPSRG30 (without the intron).

MUG assays (Jefferson et al. 1987) of the GUS enzyme activity in the established transgenic suspension lines revealed that the callus lines from construct pRESQ4 had an average of 29-fold GUS enzyme activity over the lines harboring construct pPSRG30 (Fig. 1c). A parallel histochemical assay also showed a clear difference in GUS enzyme activity between the two groups of suspension lines (Fig. 1c). The results indicated that the *rubi3* intron had a strong effect in overall gene expression enhancement in rice suspension cells.

Nuclear run-on experiments revealed nearly twofold transcriptional enhancement by the *rubi3* 5′ UTR intron

Nuclear run-on experiment is the method for testing whether a gene is regulated at the transcriptional level. During the assay, "new initiation of transcription is inhibited, and only established transcription complexes can elongate in isolated nuclei" (Eick et al. 1994). To assess whether the *rubi3* 5′ UTR intron enhances transcription, we performed nuclear run-on experiments on each of the six individual rice suspension lines transformed with the intron-containing or the intronless construct. Using the transcripts of a rice actin (*Act1*) gene as a reference makes it possible to normalize the data and to compare the *GUS* gene transcription activity among the lines. As shown in Fig. 2, the relative counts of the *GUS* transcripts in callus lines transformed with the intron-containing pRESQ4 construct were 2.9-fold higher than the ones from the lines of the intronless construct pPSRG30. The data were from two independent experiments and the difference was statistically significant ($P = 0.01849$). The results suggested an enhanced transcription by the *rubi3* 5′ UTR intron. It is possible that this increased transcription was an over-estimate considering that the length of the transcribed region in pRESQ4 (3.3 kb) was substantially longer than that of pPSRG30 (2.2 kb) as a result of the length of the intron (1.1 kb) in construct pRESQ4 in our experimental system.


Fig. 2 Transcriptional enhancement of IME in transgenic rice suspension lines. **a** Slot blot hybridization results of nuclear run-on assays from the two groups of the transgenic suspension lines. Lambda DNA or linearized plasmid DNA (5 μg each), containing the *GUS* gene or the rice *Act1* gene, was immobilized on the membrane as indicated. Each slot blot was hybridized with [α-^{32}P]-UTP labeled transcripts from a nuclear run-on reaction of the individual transgenic suspension line, as indicated. **b** Quantification of *GUS* transcripts. Radioactive signals of the *GUS* transcripts in panel **a** was quantified and normalized by the signals of the rice *Act1* reference transcripts. The mean and standard deviation of relative *GUS* transcript levels of the two groups of lines, from two replicates of each transgenic line, are presented. *Asterisks* indicates the difference is statistically significant at $P < 0.05$ level

Thus, about one third of the counts from the pRESQ4 lines could be attributed to the transcription complexes bound to the intron sequence at the time when the nuclei were isolated. However, taken the case into consideration, the intron-containing lines would still have approximately 1.9-fold higher *GUS* transgene transcription than the lines harboring the intronless pPSRG30. The results demonstrated that the *rubi3* 5′ UTR intron did enhance transcription moderately.

Estimation of *GUS* mRNA accumulation by RealTime quantitative RT-PCR

So far, IME in monocots was mostly assayed by the reporter gene enzyme activities, and mRNA accumulation affected by an intron in monocots was rarely evaluated quantitatively. In our experiments, the *GUS* mRNA accumulation in the six transgenic lines was first evaluated by northern hybridization. The intron-containing construct, pRESQ4, yielded much higher amount of the *GUS* mRNA than that from the intronless construct, pPSRG30, which was almost undetectable (Fig. 3a) and was difficult to quantify with a phosphorimager. To have a more accurate estimate of the enhancement at the mRNA level, we performed
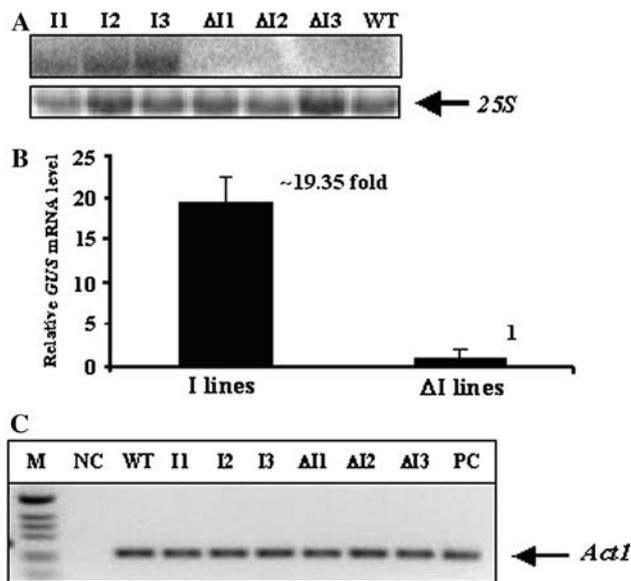

Fig. 3 *GUS* mRNA accumulation in transgenic rice suspension lines. **a** Northern analysis of *GUS* mRNA in rice suspension lines transformed with the gene construct with intron (I lines) or without the intron (ΔI lines). Rice 25S rRNA, as stained by EtBr, is also shown as a loading reference. **b** Mean and standard deviation of the relative *GUS* mRNA accumulation of the two groups of transgenic lines from quantitative RealTime RT-PCR assays. **c** Gel electrophoresis of the RT-PCR results of the rice *Act1* gene in all the lines, as a reference. The expected PCR fragment from *Act1* cDNA is indicated by an *arrow*. *M* molecular markers; *NC* negative control (no cDNA in the reaction); *WT* cDNA from non-transgenic callus; *PC* plasmid DNA control

RealTime qRT-PCR to quantify the *GUS* mRNA accumulation in these transgenic lines, normalized with the Rice *Act1* mRNA. The qRT-PCR analysis revealed an average of 19.35-fold in *GUS* mRNA accumulation in the pRESQ4 lines over the pPSRG30 lines (Fig. 3b). The results suggested that the enhanced GUS enzyme activity was mostly associated with an increased mRNA accumulation. It also indicated that the enhanced transcription only contributed a small fraction of the increased mRNA accumulation. Northern analysis revealed that the *rubi3* intron in pRESQ4 construct was well spliced in the suspension lines since no unspliced *GUS* RNA was detected. Comparing the enhancement in GUS enzyme activity (29 fold) and the mRNA level (near 20 fold), the results also implied a moderate enhancement (~45%) at the translational level by the *rubi3* intron.

GUS expression with shortened exon 1 sequence

An exon junction complex (EJC) has been shown to play an important role in the intron-meditated enhancement of gene expression in animal cells. EJC binds at 20–24 nt upstream of the splicing site after intron splicing (Le Hir et al. 2000). Genes highly homologous to those of some human EJC proteins have been identified in plants (*Arabidopsis*: Pendle et al. 2005; rice: our unpublished results). However, it remains to be seen whether a similar EJC is formed in plants and functions in IME. To gain insight in this aspect, we designed a shortened exon 1 construct SE, which was the same as that of the pRESQ4 (also called WT, or wild type, in this experiment) except that SE had only 14 bp of the exon 1 sequence (a fusion of 7 bp from each terminus of the exon 1). It was hypothesized that the SE would not have much expression enhancement by the intron if the presumed plant EJC also bound 20–24 nt upstream of the exon junction and if the EJC played a major role in plant IME. Two derivatives from these two constructs, without the 5′ UTR intron and otherwise identical to the original plasmids, pPSRG30 (also called WTΔI in this experiment) and SEΔI, were used as controls for the experiment. In a transient *GUS* expression assay, however, unexpected results were observed. First, the exon 1 sequence had a positive role in *GUS* gene expression, more or less independent from the intron. Removal of a majority of the exon 1 sequence from the intronless construct (SEΔI vs. WTΔI; Fig. 4) resulted in a sixfold reduction in GUS activity, or a threefold reduction for a similar pair of constructs that contained the intron (SE vs. WT, Fig. 4). Second, although the overall GUS activities were reduced by shortening the exon 1 sequence, it appeared that this deletion did not affect much of IME since GUS activity of SE was approximately 22-fold of the one from SEΔI, similar to the fold difference of the pair of WT and WTΔI (Fig. 4).

Fig. 4 Evaluation of IME in constructs with shortened exon 1. **a** The *GUS* gene constructs with full length 5′ UTR exon 1 (67 bp, wild type/WT/pRESQ4), or shortened 5′ UTR exon 1 (14 bp, SE), and the derivatives from the two without the 5′ UTR intron, WTΔI (pPSRG30) and SEΔI. **b** Transient GUS assay results of the four constructs from three replicates
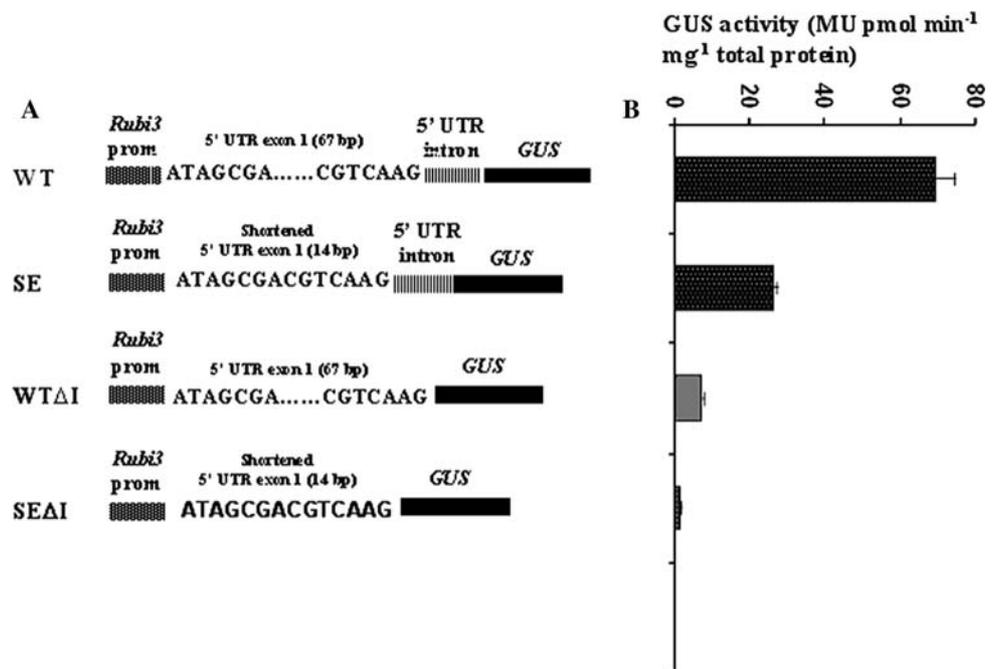
Discussion

As part of the eukaryotic gene regulatory mechanism, many introns mediate gene expression enhancement. Although the effects are rather profound in many cases studied, for quite a while the mechanism was unknown. Advances in animal cell studies reveal that IME results from a complicated process involving expression enhancement at various levels, including transcriptional and post-transcriptional (RNA accumulation and translation), and a newly identified protein complex, exon junction complex, or EJC, which binds the spliced mRNA, is mostly responsible for the expression enhancement (Le Hir et al. 2000; Wiegand et al. 2003). Although good progress has been made in IME studies in plants in recent years (Rose 2004; Belostotsky and Rose 2005; Chung et al. 2006), and some highly conserved homologs of EJC components have been identified in plants (Pendle et al. 2005), the mechanism of IME in plants and its relationship to a potential plant EJC are still not clear.

IME has been widely employed to enhance transgene expression in monocots, e.g., each of the two mostly used promoters for monocot transformation, maize *Ubi1* and rice *Act1* gene promoters, has a 5′ UTR intron (Christensen and Quail 1996; McElroy et al. 1990). In this research, we employed stably-transformed rice suspension lines containing a single-copy *GUS* reporter gene driven by the rice *rubi3* promoter, with or without its 5′ UTR intron, to quantitatively dissect the IME effects at various gene expression levels as a step to gain insight of the IME mechanism in monocot plants. Using the approach, we were able to perform nuclear run-on experiments and to detect a nearly twofold increase at the transcriptional level by the 5′ UTR

intron of the rice *rubi3* gene, which was comparable to the reported intron-mediated transcriptional enhancement in mammal cells (Furger et al. 2002). Previous studies in dicot plants failed to detect transcriptional enhancement in stably transformed tissues (Dean et al. 1989; Rose and Last 1997; Rose and Beliakoff 2000), most likely because the experimental systems used were not sensitive enough to allow such a detection as suggested by Dean et al. (1989). Our results, for the first time, experimentally demonstrated a connection between transcription and the IME by a plant intron. It was suggested that the splicing factors bound to the nascent RNA may interact with RNA Pol II C-terminal domain (CTD) and help regulate transcriptional initiation and elongation. Alternatively, the proximal intron may facilitate the release and rapid recycling of certain transcription initiation factors for new initiation events (Furger et al. 2002).

In our experiments, quantitative analysis of transcription rate and the accumulated mRNA of the *GUS* gene using the same transgenic cell lines also revealed that the enhancement at the transcriptional level was moderate, and post-transcriptional event(s) contributed the most to the observed IME. This disagrees with the suggestion that IME can be mostly attributed to transcriptional events in plants (Rose 2004). However, the mechanism of the remarkably higher accumulation of mRNA from the intron-containing construct is still poorly understood. It was reported in mammalian and plant cells that introns did not alter cytoplasmic mRNA stability (Rethmeier et al. 1997; Lu and Cullen 2003; Nott et al. 2003; Rose 2004), and the enhanced accumulation of mRNA was most likely attributed to nuclear events. Huang and Gorman (1990) observed proportionally

higher nuclear poly A+ RNA from an intron-containing construct in mammalian cells and suggested that splicing was coupled with the polyadenlyation/transport pathway. Lu and Cullen (2003) revealed that a major effect of introns in mammalian cells was to enhance 3′ processing of RNA. Nott et al. (2003) also observed the effect of an intron on RNA polyadenylation, which may facilitate stability and export of the transcribed RNA. No such data are available in plants at the moment. However, considering that the two constructs (intron-containing and intronless) used in this report differed at transcriptional initiation rate by only two-fold, transcription from the intronless construct was still quite strong (Fig. 2). Most of the transcripts from this construct must have been rapidly degraded before they reached the cytoplasm. Bousquet-Antonelli et al. (2000) revealed a nuclear pathway that "rapidly degrades unspliced pre-mRNAs in yeast", which mainly involved the exosome complex, and exosome-like complex has been reported in plants (Chekanova et al. 2002). It would be interesting to see whether exosome has a role in the degradation of the pre-mRNA from an intronless construct in plants.

It has been demonstrated that EJC, rather than the splicing itself, is mainly responsible for the intron-mediated enhancement in mammalian cells (Wiegand et al. 2003; Nott et al. 2004). In our quest for a relationship between IME and the potential EJC in plants, we designed gene constructs (SE and SEΔI) with the exon 1 shortened to 14 bp. The hypothesis was that the exon 1 would be too short to provide a binding site for EJC on the spliced RNA and thus would abort the IME if the proposed plant EJC bound similarly and was mainly responsible for the observed IME as in mammals. However, to our surprise, the SE construct, which had the intron, still had similar enhancement (22-fold) when compared with SEΔI, which was intronless and would yield identical mRNA. The results may suggest a divergence in the mechanisms of IME between plants and animals: plants either do not have the equivalent of the EJC of animal cells, or the plant EJC does not bind that far (20–24 nt upstream), or the plant EJC does not play a similar role in IME. In line with the notion, Pendle et al. (2005) reported the *Arabidopsis* EJC homologues were located in the nucleolus whereas their counterparts in mammalian cells were predominantly located in the nucleoplasm and appeared to be excluded from the nucleolus (Kataoka et al. 2000, 2001).

Results shown in Fig. 4 also indicated an enhancing role of the untranslated exon 1 of the *rubi3* gene in gene expression in the presence or absence of the 5′ UTR intron. Although 5′ untranslated leader sequences from certain plant RNA viruses have been intensively studied in elevating translation in plant cells and used in enhancing transgene expression (Gallie 1993), the 5′ untranslated exon sequences from plant genes were less studied. Clancy et al.

(1994) observed that the untranslated exon 1 sequence from maize *Shrunken1* gene enhanced CAT reporter gene expression by 7.3-fold when the 5′ UTR intron was removed from the construct. It remains to be determined at what expression level(s) the exon-mediated enhancement takes place.

It became clear in the past several years that translation contributed to the IME in mammalian cells. Core components of the EJC, such as Magoh and Y14, and a few splicing factors, remained stably associated with the spliced mRNA after nuclear export, and were likely to have a role in promoting polysome association with the spliced mRNA to improve translational efficiency (Nott et al. 2004; Sanford et al. 2004). In plants, data also implied improved translation by introns (Bourdon et al. 2001; Rose 2004). Our results indicated approximately 45% enhancement by the *rubi3* 5′ UTR intron at the translational level. Since it built upon the enhancement at previous expression levels, the overall IME was further elevated from 20-fold (mRNA accumulation level) up to 29-fold (at protein level as measured by GUS enzyme activity).

In summary, we quantitatively dissected IME by the 5′ UTR intron of the rice *rubi3* gene at various gene regulation levels in transgenic rice suspension cells, and revealed a nearly twofold increase at the transcriptional level, about 20-fold augment in mRNA accumulation, and 29-fold enhancement at the protein level, suggesting moderate IME at both transcriptional and translational levels, and a remarkable enhancement at the post-transcriptional level. The results also indicated a positive role (three to sixfold) of the exon 1 in gene expression regulation, and hinted that the mechanisms of IME in plants and animals might be diverse.

References

Azhakanandam K, McCabe MS, Power JB, Lowe KC, Cocking EC, Davey MR (2000) T-DNA transfer, integration, expression and inheritance in rice: effects of plant genotype and *Agrobacterium* super-virulence. J Plant Physiol 157:429–439

Belostotsky DA, Rose AB (2005) Plant gene expression in the age of systems biology: integrating transcriptional and post-transcriptional events. Trends Plant Sci 10:347–353

Blockland RV, Geest NV, Mol JNM, Kooter JM (1994) Transgene-mediated suppression of chalcone synthase expression in *Petunia hybrida* results from an increase RNA turnover. Plant J 6:861–877

Bourdon V, Harvey A, Lonsdale DM (2001) Introns and their positions affect the translational activity of mRNA in plant cells. EMBO Rep 2:394–398

Bousquet-Antonelli C, Presutti C, Tollervey D (2000) Identification of a regulated pathway for nuclear pre-mRNA turnover. Cell 102:765–775

Buchman AR, Berg P (1988) Comparison of intron-dependent and intron-independent gene expression. Mol Cell Biol 8:4395–4405

Callis J, Fromm ME, Walbot V (1987) Introns increase gene expression in cultured maize cells. Genes Dev 1:1183–1200

Chekanova JA, Dutko JA, Mian IS, Belostotsky DA (2002) *Arabidopsis thaliana* exosome subuint AtRrp4p is a hydrolitic $3' \rightarrow 5'$ exonuclease containing S1 and KH RNA binding domains. Nucleic Acids Res 30:695–700

Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res 5:213–218

Chung BYW, Simons C, Firth AE, Brown CM, Hellens RP (2006) Effect of 5′UTR introns on gene expression in *Arabidopsis thaliana*. BMC Genomics 7:120

Clancy M, Hannah LC (2002) Splicing of the maize *Sh1* first intron is essential for enhancement of gene expression, and a T-rich motif increases expression without affecting splicing. Plant Physiol 130:918–929

Clancy M, Vasil V, Hannah LC, Vasil IK (1994) Maize *Shrunken-1* intron and exon regions increase gene expression in maize protoplasts. Plant Sci 98:151–161

Dean C, Favreau M, Bondnutter D, Bedbrook J, Dunsmuir P (1989) Sequences downstream of translation start regulate quantitative expression of 2 *Petunia rbcs* genes. Plant Cell 1:201–208

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 14:19–21

Egoavil C, Marton HA, Baynton CE, McCullough AJ, Schuler MA (1997) Structural analysis of elements contributing to 5′ splice site selection in plant pre-mRNA transcripts. Plant J 12:971–980

Eick D, Kohlhuber LF, Wolf DA, Strobl LJ (1994) Activation of pausing RNA polymerases by nuclear run-on experiments. Anal Biochem 218:347–351

Furger A, O'Sullivan JM, Binnie A, Lee BA, Proudfoot NJ (2002) Promoter proximal splice sites enhance transcription. Genes Dev 16:2792–2799

Gallagher SR (1992) GUS protocols: using the GUS gene as a reporter of gene expression quantitation of GUS activity by fluorometry. Academic press Inc., New York

Gallie DR (1993) Posttranscriptional regulation of gene-expression in plants. Annu Rev Plant Physiol Plant Mol Biol 44:77–105

Goodall GJ, Filipowicz W (1989) The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell 58:473–483

Goodall GJ, Kiss T, Filipowicz W (1991) Nuclear RNA splicing and small nuclear RNAs and their genes in higher plants. Oxf Surv Plant Mol Cell Biol 7:255–296

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J 6:271–282

Huang MTF, Gorman CM (1990) Intervening sequences increase efficiency of RNA 3′ processing and accumulation of cytoplasmic RNA. Nucleic Acids Res 18:937–947

Ingelbrecht I, de Carvalho F (1992) Isolation of nuclei and in vivo run-off transcription. In: Inze D, Van der Straeten D, Van Montagu M (eds) EMBO practical course on plant molecular biology. Laboratorium voor Genetica, Gent, pp 117–132

Ingham DJ, Beer S, Money S, Hansen G (2001) Quantitative real-time PCR assay for determining transgene copy number in transformed plants. Biotechniques 31:132–140

Initiative AG (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Jefferson RA, Kavanagh TA, Bevan MW (1987) Gus fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6:3901–3907

Kataoka N, Yong J, Kim VN, Velazquez F, Perkinson RA, Wang F, Dreyfuss G (2000) Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. Mol Cell 6:673–682

Kataoka N, Diem MD, Kim VN, Yong J, Dreyfuss G (2001) Magoh, a human homolog of Drosophila mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. EMBO J 20:6424–6433

Kim BR, Nam HY, Kim SU, Kim SI, Chang YJ (2003) Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. Biotechnol Lett 25:1869–1872

Klein TM, Gradziel T, Fromm ME, Sanford JC (1988) Factors influencing gene delivery into *Zea mays* cells by high-velocity micro-projectiles. Biotechnology 6:559–563

Lambermon MHL, Simpson GG, Kirk DAW, Hemmings-Mieszczak M, Klahre U, Filipowicz W (2000) UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. EMBO J 19:1638–1649

Le Hir H, Izaurralde E, Maquat LE, Moore MJ (2000) The spliceosome depositsmultiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. EMBO J 19:6860–6869

Lu J (2006) Gene expression regulation by the 5′ regulatory sequences of the rice *rubi3* gene in transgenic rice plants. Ph.D. thesis. North Carolina State University. http://www.lib.ncsu.edu/theses/available/etd-04272006-145216/unrestricted/etd.pdf

Lu SH, Cullen BR (2003) Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. RNA 9:618–630

Luehrsen KR, Walbot V (1991) Intron enhancement of gene-expression and the splicing efficiency of introns in maize cells. Mol Gen Genet 225:81–93

Luehrsen KR, Walbot V (1994) Addition of A-rich and U-rich sequence increases the splicing efficiency of a deleted form of a maize intron. Plant Mol Biol 24:449–463

Luehrsen KR, Taha S, Walbot V (1994) Nuclear pre-messenger RNA processing in higher plants. Nucleic Acid Res 47:149–193

Mascarenhas D, Mettler IJ, Pierce DA, Lowe HW (1990) Intron-mediated enhancement of heterologous gene expression in maize. Plant Mol Biol 15:913–920

McCullough AJ, Schuler MA (1997) Intronic and exonic sequences modulate 5' splice site selection in plant nuclei. Nucleic Acids Res 25:1071–1077

McElroy D, Zhang WG, Cao J, Wu R (1990) Isolation of an efficient actin promoter for use in rice transformation. Plant Cell 2:163–171

Nott A, Muslin SH, Moore MJ (2003) A quantitative analysis of intron effects on mammalian gene expression. RNA 9:607–617

Nott A, Le Hir H, Moore MJ (2004) Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Gene Dev 18:210–222

Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A (1993) Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. Genetics 135:385–404

Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, Andersen J, Mann M, Lamond AI, Brown JWS, Shaw PJ (2005) Proteomic

analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. Mol Biol Cell 16:260–269

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating rnRNA processing with transcription. Cell 108:501–512

Reddy ASN (2001) Nuclear pre-mRNA splicing in plants. Crit Rev Plant Sci 20:523–571

Rethmeier N, Seurinck J, VanMontagu M, Cornelissen M (1997) Intron-mediated enhancement of transgene expression in maize is a nuclear, gene-dependent process. Plant J 12:895–899

Rethmeier N, Kramer E, van Montagu M, Cornelissen M (1998) Identification of *cat* sequences required for intron-dependent gene expression in maize cells. Plant J 13:833–835

Rose AB (2002) Requirements for intron-mediated enhancement of gene expression in *Arabidopsis*. RNA 8:1444–1453

Rose AB (2004) The effect of intron location on intron-mediated enhancement of gene expression in *Arabidopsis*. Plant J 40:744–751

Rose AB, Beliakoff JA (2000) Intron-mediated enhancement of gene expression independent of unique intron sequence and splicing. Plant Physiol 122:535–542

Rose AB, Last RL (1997) Introns act post-transcriptionally to increase expression of the *Arabidopsis thaliana* tryptophan pathway gene PAT1. Plant J 11:455–464

Sanford JR, Gray NK, Beckmann K, Caceres JF (2004) A novel role for shuttling SR proteins in mRNA translation. Genes Dev 18:755–768

Simpson GG, Filipowicz W (1996) Splicing of precursors to mRNA in higher plants: Mechanism, regulation and sub-nuclear organisation of the spliceosomal machinery. Plant Mol Biol 32:1–41

Simpson CG, Simpson GG, Clark G, Leader DJ, Vaux P, Guerineau F, Waugh R, Brown JWS (1992) Splicing of plant pre-messenger RNAs. Proc R Soc Edinb B Biol Sci 99:31–50

Sivamani E, Qu R (2006) Expression enhancement of a rice polyubiquitin gene promoter. Plant Mol Biol 60:225–239

Sivamani E, Huet H, Shen P, Ong CA, de Kochko A, Fauquet C, Beachy RN (1999) Rice plant (*Oryza sativa* L.) containing Rice tungro spherical virus (RTSV) coat protein transgenes are resistant to virus infection. Mol Breed 5:177–185

Wiegand HL, Lu SH, Cullen BR (2003) Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. Proc Natl Acad Sci USA 100:11327–11332

Yon J, Fried M (1989) Precise gene fusion by PCR. Nucleic Acids Res 17:4895–4895

Zhang S, Song W, Chen L, Ruan D, Taylor N, Ronald P, Beachy RN, Fauquet CM (1998) Transgenic elite *Indica* rice varieties, resistant to *Xanthomonas oryzae* pv. *Oryzae*. Mol Breed 4:551–558