Journal of Integrative Agriculture 2016, 15(8): 1716–1726

Available online at www.sciencedirect.com

ScienceDirect

RESEARCH ARTICLE

Utilizing modified *ubi1* introns to enhance exogenous gene expression in maize (*Zea mays* L.) and rice (*Oryza sativa* L.)

PAN Yang-yang[1*], CHEN Rui[2*], ZHU Li[1], WANG Hai[1], HUANG Da-fang[1], LANG Zhi-hong[1]

[1] *Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China*

[2] *Tianjin Institute of Agricultural Quality Standard and Testing Technology, Tianjin Academy of Agricultural Sciences, Tianjin 300192, P.R.China*

Abstract

The phenomenon of intron-mediated enhancement (IME) was discovered in 1990 based on the observation that plant introns can stimulate gene expression, particularly in monocots. However, the intrinsic mechanism of IME remains unclear because many studies have yielded various results depending on the promoter, reporter gene, flanking sequences of the intron, and target cell or tissue. In this study, the effect of the first intron of the maize ubiquitin gene (*ubi1* intron) was investigated by changing insertion sites, deleting specific regions and mutating individual motifs in maize (*Zea mays* L.) and rice (*Oryza sativa* L.) using *ubi1* intron-containing *GUS* (*β-glucuronidase*) constructs. In maize callus, the integration of the full-length *ubi1* intron into the *GUS* coding sequence at the +13, +115 and +513 positions by particle bombardment increased GUS activity approximately five-, four- and two-fold, respectively. Eight truncated *ubi1* introns in the pSG(13i)N construct significantly influenced *GUS* gene expression to different degrees in transient assays. Notably, the 3′ region deletions significantly reduced the IME effect, whereas a 142-nt deletion, pSG(13i-P5)N, in the 5′ region caused a 1.5-fold enhancement relative to pSG(13i)N. Furthermore, four site-directed mutageneses were performed in pSG(13i-P5)N; these constructs resulted in the up-regulation of *GUS* gene expression to different levels. The most effective modified *ubi1* intron, pSG(13i-M4)N, was further evaluated and proved in rice using transient experiments. In addition, the sequences flanking the *GUS* insertion significantly influenced the IME effect of the vectors that were constructed. The modified *ubi1* intron had the potential application on crop genetic engineering.

Keywords: maize, intron-mediated enhancement, *ubi1* intron, intron modification, IME signals

1. Introduction

Introns are intervening sequences of coding genes in eukaryotic cells that can be removed from pre-mRNAs by a two-step successive in-line trans-esterification reaction (Moore and Sharp 1993). Callis *et al*. (1987) demonstrated that the first intron of alcohol dehydrogenase 1 (*Adh1*) in Black Mexican Sweet Corn protoplasts could improve chloramphenicol acetyl transferase gene (CAT) activity. Many subsequent studies revealed that introns could strengthen

Received 10 August, 2015 Accepted 8 December, 2015
PAN Yang-yang, Mobile: +86-15210675507, E-mail: yypan@genetics.ac.cn; CHEN Rui, Mobile: +86-15900351608, E-mail: chenrui.taas@gmail.com; Correspondence LANG Zhi-hong, Tel: +86-10-82109857, Fax: +86-10-82106142, E-mail: langzhihong@caas.cn
* These authors contributed equally to this study.

doi: 10.1016/S2095-3119(15)61260-6

gene expression in eukaryotes, including vertebrates, invertebrates, fungi, and plants (Buchman and Berg 1988; Huang and Gorman 1990; Choi *et al*. 1991; Meredith and Storti 1993; Okkema *et al*. 1993). This positive effect of introns has been called intron-mediated enhancement (IME) (Mascarenhas *et al*. 1990). Thus far, the underlying mechanism of IME remains unclear, although several studies have identified the potential common features or important characteristics of IME. First, the enhancing introns are directional, in contrast to enhancers in eukaryotic gene expression (Rose 2004). Second, the intron location, flanking sequences and internal sequences are important for the performance of IME (Le *et al*. 2003). Importantly, only some introns are able to enhance gene expression, including lead introns and promoter-proximal introns. Generally, the enhancing introns must be relatively conserved to stimulate gene expression (Clancy *et al*. 1994; Rose and Beliakoff 2000). Not all internal sequences of introns can play the same role, and in some cases, parts of the same sequence are redundant. Clancy and Hannah (2002) reported that only 14% of the *Sh1* first intron could produce a full enhancement of gene expression. Third, introns have global regulatory functions throughout the entire process of gene expression. Introns enhance gene activation in the nucleus and promote transcription initiation, mRNA export from the nucleus, mRNA stability, translation efficiency, and non-mediated mRNA degradation (NMD) (Chasin 2007; Moore and Proudfoot 2009). Niu and Yang (2011) presented the hypothesis that IME of gene expression could reduce transcription-associated genome instability by inhibiting topoisomerase I cutting activity. The latest study demonstrated that the intron-containing gene forms a gene loop, wherein the promoter region interacts with the 5′ splicing site and the terminator interacts with the 3′ splicing site, resulting in increased nascent transcript levels; thus, the authors proposed that this gene looping plays a role in the IME of transcription in yeast (Moabbi *et al*. 2012).

The maize ubiquitin promoter and first intron were first studied in monocotyledons by Bruce (1989). Subsequently, Vain found that the *ubi1* intron with a 41-nt exonic flanking sequence (p35SubiGUS) is more effective at enhancing gene expression than other carriers in suspended maize cells (Vain *et al*. 1996), which suggests that the modification of the *ubi1* intron has the potential to regulate gene expression.

In 2008, Rose *et al*. (2008) proposed an algorithm called IMEter, which is able to predict the IME potential of introns *via* bioinformatics analysis; the enhancing motifs were speculated to be GATCTG in *Arabidopsis thaliana* and TCGATC in rice (*Oryza sativa* L.) (Farquharson 2008). Morello *et al*. (2010) demonstrated that the IMEter scores correlate with the GUS expression level in transiently transformed rice calli. Further studies showed that the IME signals are preferentially located at the 5′ end of first introns and in coding sequences near the transcription start site (Parra *et al*. 2011). The IME signals could influence gene expression by insertion or removal, so it maybe is an alternative way to increase the IME effect of the *ubi1* intron by increasing the number of enhancing motifs.

Herein, we evaluated the effects of modified *ubi1* intron sequences using transient gene expression in the embryogenic calli of maize to study the effect of the first *ubi1* intron on *GUS* gene expression in maize. The introns were inserted into the *GUS* gene coding sequences near to the initiation codon AUG, which is close to the eukaryotic genetic structure. This approach may contribute to the correct splicing of *GUS* pre-mRNA and effectively prevent gene escape to prokaryotic organisms; the possibility of gene escape has increased the biosafety issues in transgenic research. Our results indicated that different locations, dissimilar sequences of the *ubi1* intron and the abundance of the enhancing motif could generate different effects. Two modified introns with high IME efficiency, pSG(13i-P5)N and pSG(13i-M4)N, were generated that may be used for the stable expression of exogenous genes.

2. Materials and methods

2.1. Maize and rice tissue cultures

Embryogenic callus cultures of maize (*Zea mays* L. hybrid line HiII and inbred Zong 31) were prepared for microprojectile bombardment transformation. The immature maize embryos were isolated from kernels after 10–15 days of self-pollination and cultured on modified N6 medium (2,4-D, 2 mg L^{-1}; L-proline, 690 mg L^{-1}, pH 5.8) to induce callus formation. Type II embryonic calli were selected for subculture at an interval of 15 d (Bruce 1989). The rice calli (*Oryza sativa cv*. Nipponbare) were derived from mature seeds according to standard procedures (Samadder *et al*. 2008). The calli were subcultured at an interval of 15 d.

2.2. Plasmid construction

The constructs used in this work are shown in Figs. 1-A, 2-A, 3-A, 4-A, 5-A, and 6-A. Plasmid pSGN contains the CaMV35S promoter, *GUS* coding sequences and nopaline synthase terminator; this construct is deposited in the laboratory. Three locations in the *GUS* coding sequences were selected to create *Pst*I enzyme sites (CTGCAG) by site-directed mutagenesis; these sequences are at positions +13 to +18, +115 to +120 and +513 to +518. The full-length *ubi1* intron was inserted as a *Pst*I fragment into these three *Pst*I sites in the *GUS* coding sequence to generate the vectors

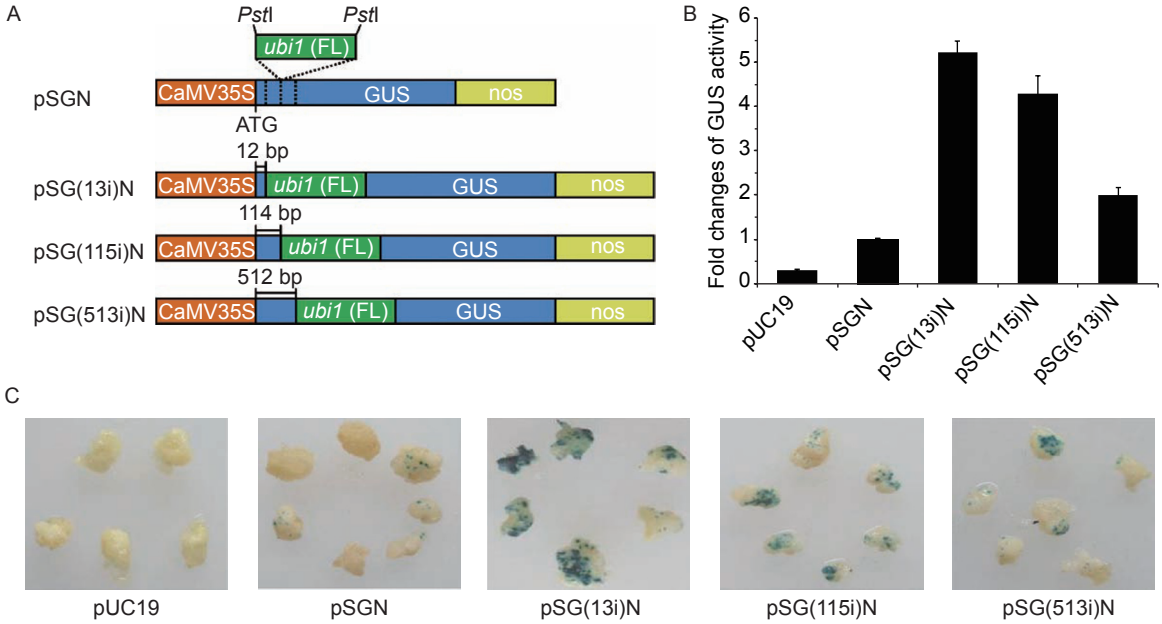

Fig. 1 Intron-mediated enhancement (IME) testing of the full-length *ubi1* intron at different positions in the *GUS* gene. Diagrams of the tested constructs with the full-length *ubi1* intron cassette. The three insertion sites, +13, +115 and +513, are indicated by vertical dotted lines in the *GUS* coding sequences. The +13 to +18, +115 to +120 and +513 to +518 sequences of the *GUS* gene were independently mutated to *Pst*I restriction enzyme sites, and the full-length *ubi1* intron was inserted *via Pst*I to generate pSG(13i)N, pSG(115i)N and pSG(513i)N (A). Fluorescence measurements of GUS enzymatic activity in transformed maize calli bombarded by different vectors, and the pSGN construct was used as a control (B). GUS histochemical staining of transformed maize calli bombarded by different vectors (C). Bars represent SE (standard error). The same as below.

pSG(13i)N, pSG(115i)N and pSG(513i)N.

To investigate the function of truncated *ubi1* introns, the *GUS* coding sequences were scanned for the two-nucleotide sequences GT (in the 5′ portion) and AG (in the 3′ portion), and a series of predicted truncated GT-AG *ubi1* introns (Table 1) were chosen to evaluate the IMEter scores as described by Rose *et al*. (2008). The truncated fragments P1, P2 and P3 resulted from the deletion of the 3′ sequence of the *ubi1* intron to nucleotides +640, +714 and +806, respectively, and the truncated fragments P4, P5 and P6 resulted from the deletion of 5′ upstream regions to nucleotides +51, +142 and +280 sites of the *ubi1* intron. Two fragments removing nucleotides +52 to +142 and +714 to +806 from the *ubi1* intron produced A and B, respectively. The above eight truncated *ubi1* introns replaced the *ubi1* intron in vector pSG(13i)N as *Pst*I fragments to generate eight constructs (Fig. 3-A).

For the mutation of motifs in the *ubi1* intron, four derivatives were produced based on sequence P5 in Fig. 4-A: fragments M1, M2, M3, and M4 were derived from the mutation of the six nucleotides at positions +11 to +16, +78 to +83, +282 to +287, and +510 to +515 to TCGACT. Then, the four modified *ubi1* introns were inserted into the pSG(13i)N vector as *Pst*I fragments to replace the full-length *ubi1* intron. The IMEter scores of the mutated *ubi1* introns are listed in Table 1.

For universal vector design, the *ubi1* intron, *ubi1*-P5 intron and *ubi1*-M4 intron were inserted immediately after the initiator AUG codon or at the +13 site of the *GUS* coding sequences as *Pst*I–*Sal*I fragments to generate six simple universal vectors referred to as pSAiGN (full-length *ubi1*), pSAP5GN, pSAM4GN, pSA12iGN, pSA12P5GN, and pSA12M4GN (Fig. 6-A and B). All constructed vectors in this study were entirely sequenced for confirmation. The plasmids were isolated and purified from *Escherichia coli* for transformation.

2.3. Particle bombardment transformation

Embryogenic callus cultures from maize and rice were prepared for microprojectile bombardment transformation. Embryogenic callus was placed on hypertonic medium (N6 agarose supplemented with 2,4-D and mannitol) for 4 h before bombardment. Bombardment was performed at 1 100 psi helium pressure and at a 7-cm distance with a PDS-1000/He (Bio-Rad, Hercules, CA, USA). The transformation by microprojectile bombardment was performed essentially as described by Zhang *et al*. (1999). Each vector was designed and confirmed 10 biological replicates, and each experiment was repeated three times.

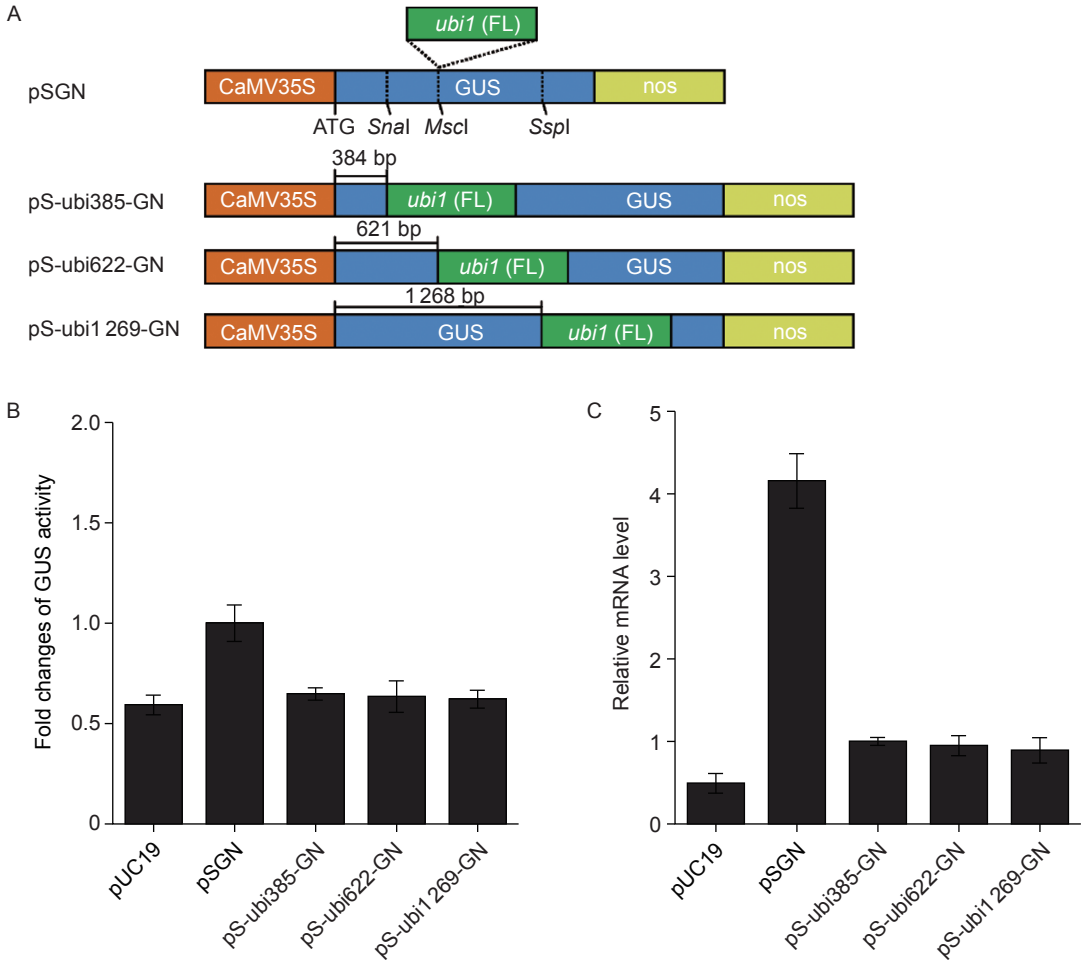

Fig. 2 IME effect of the full-length *ubi1* intron with blunt-end insertion at various locations in the *GUS* coding sequences. Diagrams of tested constructs with the full-length *ubi1* intron cassette. Three restriction enzyme sites, *Sna*BI, *Msc*I and *Ssp*I, at positions +385, +622 and +1 269 of *GUS* coding sequence respectively, were used for *ubi1* intron integration (A). Fluorescence measurements of GUS enzymatic activity in maize calli bombarded by tested constructs, and the pSGN construct was used as a control (B). qRT-PCR analysis of *GUS* transcript level in maize calli bombarded by tested constructs (C).

2.4. Histochemical and fluorometric analyses of GUS activity

At 12 h after bombardment, the bombarded calli were removed to Chu's N-6 (N6) medium. A part of each callus was selected randomly for histochemical assays after 12 h. GUS enzyme activity was measured using the substrate 4-methylumbelliferyl-β-D-glucuronide (MUG) (Jefferson *et al*. 1987). Protein content was determined with a dye-binding assay using bull serum albumin (BSA) as a standard (Bio-Rad, USA).

2.5. qRT-PCR analysis of the *GUS* cDNA

DNase-treated RNA for quantitive RT-PCR experiments was isolated from maize calli 48 h after transformation using the Invitrogen RNAspin Mini Kit (Invitrogen, USA). Reverse transcription was performed with 1 μg RNA, using oligo(dT) primer and the Invitrogen reverse transcriptase. The expression level of *GUS* mRNA in bombarded calli developed for plasmids pSGN, pS-ubi385-GN, pS-ubi622-GN, pS-ubi1269-GN, pS-ubiH-GN, and negative control pUC19 were evaluated by qRT-PCR. A plasmid containing the full-length *GUS* cDNA was used as an external control. Serial dilutions of the control plasmid were used to generate a standard curve. A housekeeping gene (maize-actin) was used as an internal control to correct any variation in samples. PCR amplification for continuous fluorescence detection was performed (Corbett Research Rotor-Gene 6000) in a total volume of 20 μL containing 1/10 μL of cDNA by using Qiagen SYBR Green PCR Kit. Each PCR reaction was performed triplicately. Quantity of the target samples was calculated by using the Rotor-gene 6 software. Melting curve analysis was done to characterize the amplified

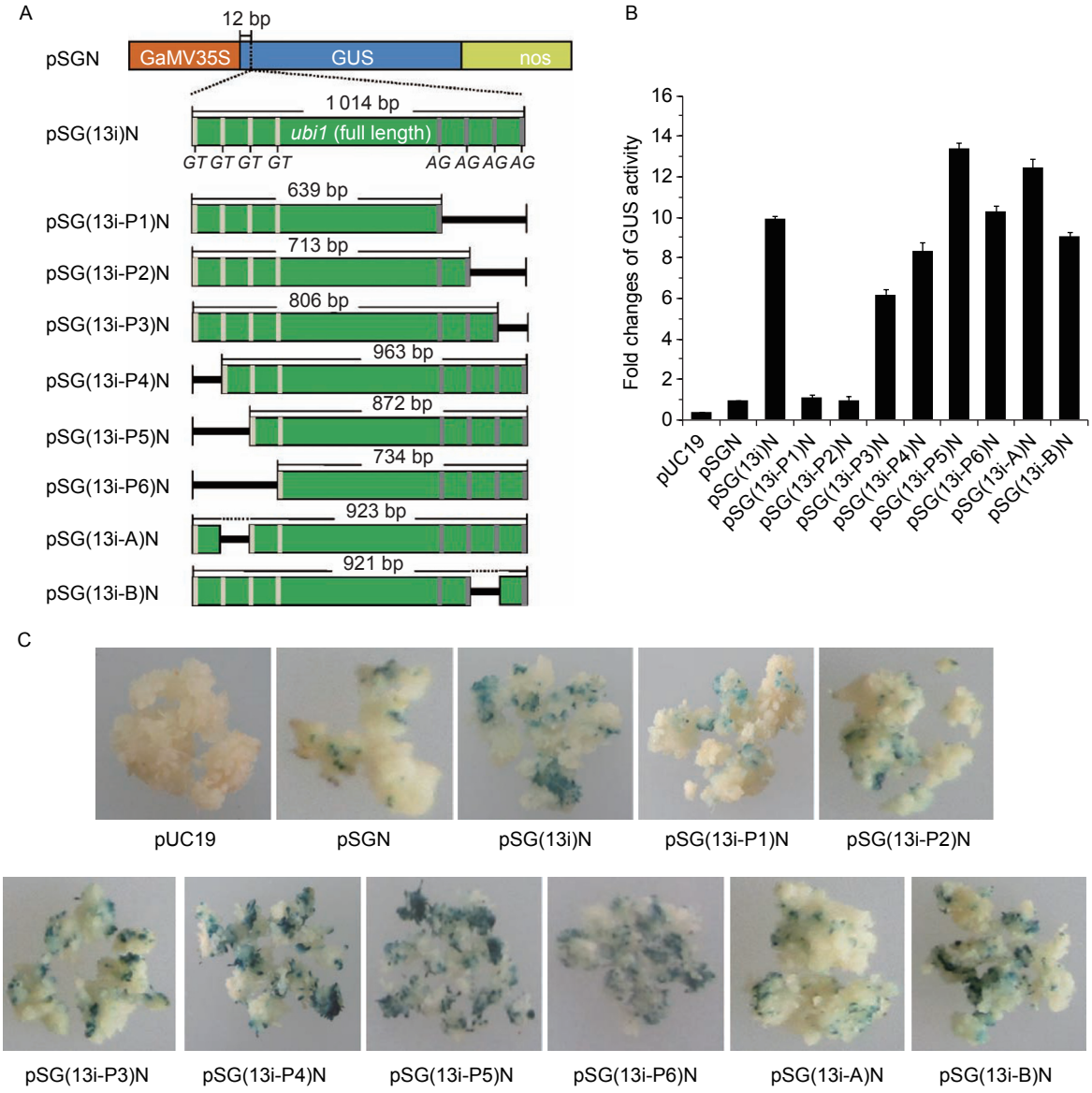

Fig. 3 IME effect of partial *ubi1* sequences based on the pSG(13i)N construct. Diagrams of the tested constructs with truncated *ubi1* intron sequences. Eight 'GT' (light grey lines) and 'AG' (dark grey lines) sites in the *GUS* coding sequences were used as coordinates for *ubi1* derivative generation and eight tested constructs includes 5′ deletion constructs, 3′ deletion constructs and internal deletion constructs. The black bars represent the deleted sequences (A). Fluorescence measurements of GUS enzymatic activity in transformed maize calli bombarded by eight tested constructs. The pUC19, pSGN and pSG(13i)N constructs were used as controls (B). The results of GUS histochemical staining assays of the maize calli transformed with the different constructs (C).

products by slowly raising the temperature from 65 to 95°C with fluorescence data collected at 0.2°C intervals. Control without the nucleic acids was included for each run.

3. Results and discussion

3.1. The full-length ubi1 intron displays an IME effect when inserted into the coding sequence of the GUS gene

The full-length *ubi1* intron was inserted as a *Pst*I fragment into the *Pst*I site by site-mutagenesis at positions +13, +115 and +513 of the *GUS* gene (relative to the translation start site; Fig. 1-A). The *ubi1* intron was cloned into the *Pst*I site without introducing any extraneous sequences (Simpson and Filipowicz 1997; Rose and Beliakoff 2000). Although the modification of the +13 and +513 sites changed one or two nucleotides, these changes did not affect the activity of the *GUS* gene.

Maize calli bombarded with pSG(13i)N, pSG(115i)N and pSG(513i)N produced many blue spots that were distributed

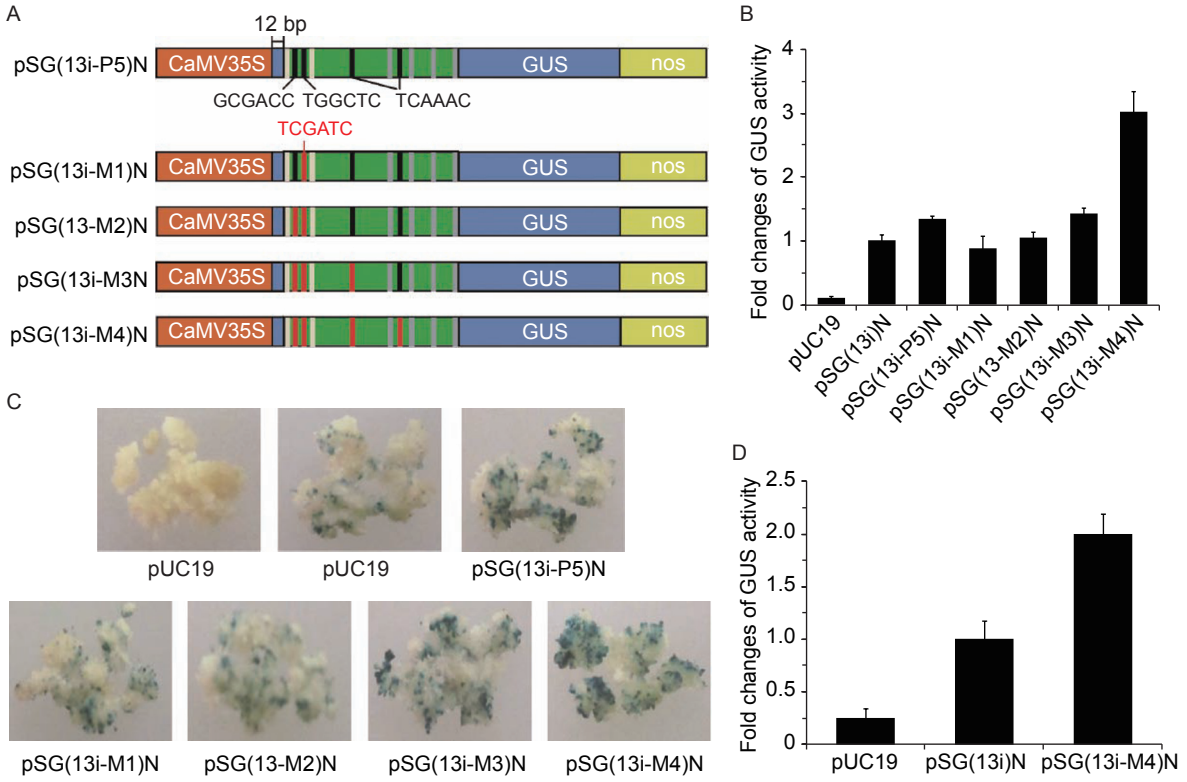

Fig. 4 Evaluation of four mutageneses based on the truncated *ubi1*-P5 intron. Diagrams of the mutation sites in the *ubi1*-P5 intron sequence. The black bars represent original elements. The mutated motif is indicated with a red bar (TCGATC) (A). Fluorescence measurements of GUS enzymatic activity in transformed maize calli and the pSG(13i)N construct was used as a control (B). GUS histochemical staining assays of maize calli bombarded by mutanted *ubi1*-P5 introns (C). Fluorescence measurements of GUS enzymatic activity in transformed rice calli transformed by *ubi1*-M4 introns. The pSG(13i)N was used as a control. The pSG(13i-M4)N showed better performance than pSG(13i)N (D).

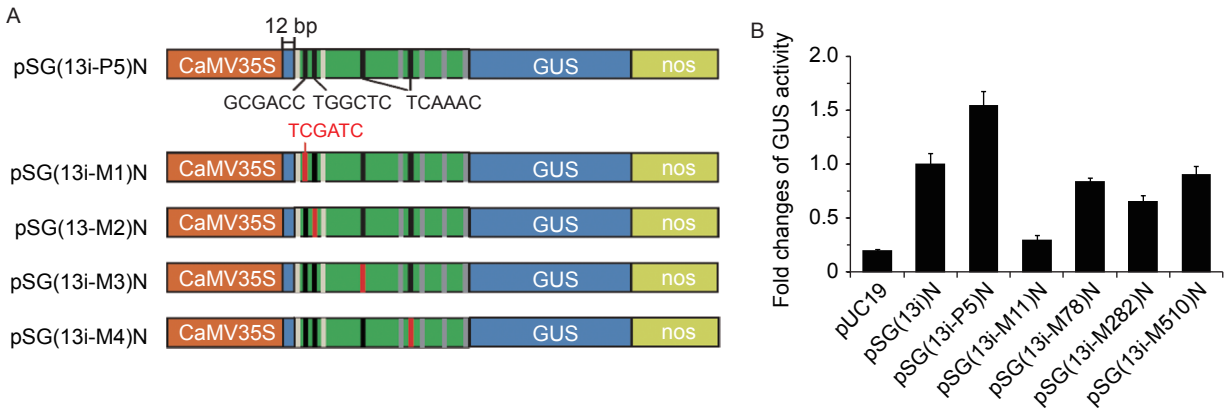

Fig. 5 IME testing of single site-directed mutagenesis of the pSG(13i-P5)N construct. Diagrams of the mutation sites in pSG(13i-P5)N. The black bars represent the original elements. The mutated motif is indicated with a red bar (TCGATC) (A). Fluorescence measurements of GUS enzymatic activity in maize calli bombarded by single site mutagenesis constructs. The pSG(13i)N construct was used as a control (B).

randomly throughout the tissues (Fig. 1-C). In some tissues, the blue spots were extremely large and dark, which suggested stronger GUS activity. The GUS staining intensity of the pSG(13i)N transfection calli was greater than those of pSG(115i)N and pSG(513i)N. Only small areas of GUS activity were detected in the explants bombarded with pSGN,

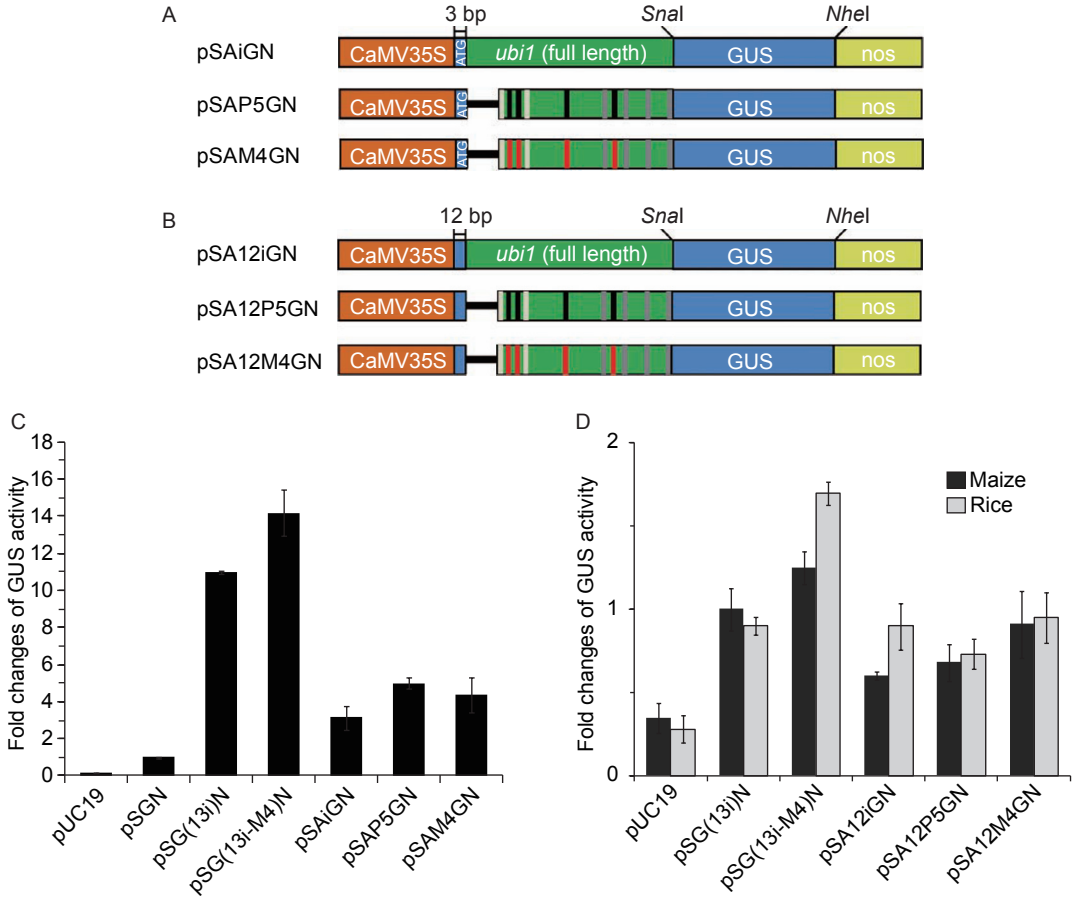

Fig. 6 Construction of universal vectors and comparative analysis of IME effects. Schematic representation of universal vectors construction adjacent to the initiation codon ATG (+4 insertion site). *Sal*I and *Nhe*I restriction sites were introduced for exogenous gene insertion (A). Schematic representation of universal vector construction with a +13 insertion site in the *GUS* gene (B). Comparative analysis of GUS enzymatic activity in maize calli bombarded by constructs of the +4 and +13 insertion sites in *GUS*. The pSGN construct was used as a control (C). GUS enzymatic activity analysis in transformed calli of maize and rice for testing the optimal universal vectors. The pSG(13i)N construct was used as a control (D).

Table 1 IMEter scores of individual *ubi1* introns and derivatives used in this study[1)]

Seq_identifier	Ver. 1.0	Ver. 2.1	Percentile
ubi1_full_length_1 014 bp	31.1	25.6	98
ubi1_P1_639	44.81	25.27	98
ubi1_P2_713	42.48	25.46	98
ubi1_P3_806	39.56	25.6	98
ubi1_P4_963	16.7	14.9	93
ubi1_P5_872	7.49	10.28	88
ubi1_P6_734	–6.04	6.4	79
ubi1_A_923	19.94	19.89	96
ubi1_B_921	33.2	25.46	98
ubi1_P5_M1	8.88	12.35	91
ubi1_P5_M2	10.41	15.6	94
ubi1_P5_M3	13.22	16.46	94
ubi1_P5_M4	16.03	16.86	95

[1)] These scores were calculated using IMEter online (http://korflab.ucdavis.edu/cgi-bin/IMEter_2014/web-imeter2.1.pl).

and no blue spots were observed in calli bombarded with pUC19, which served as an experimental negative control.

The GUS enzyme activity of different constructs containing the *ubi1* intron is shown in Fig. 1-B. The *ubi1* intron at all three positions significantly increased the GUS activity in the bombarded calli compared to the intronless construct pSGN, in which the enhancement effect decreased gradually with the distance from the initial AUG codon. The *ubi1* intron at position +13 (pSG(13i)N) significantly enhanced GUS activity to a level that was approximately 5-fold higher than that of pSGN, and the pSG(115i)N construct enhanced GUS activity to a level that was approximately 4-fold higher than that of pSGN. When the *ubi1* intron was inserted at position +513 (pSG(513i)N), its IME effect decreased to approximately 2-fold higher than that of pSGN. Further studies indicated that introns with inserted positions at +13, +115 and +513 were all correctly spliced at the 5′ and 3′

splice sites (5′-GT-AG-3′) as predicted. The effects of the intron position on gene expression have been extensively reported in the literature; the promoter-proximal introns showed a stronger enhancement of gene expression than later introns, regardless of whether the experiment used a transient transformation system or a stable transformation system (Rose 2004). However, when the *ubi1* intron was inserted at positions +385, +622 and +1 269 using the blunt-end enzyme sites for *Sna*BI (TAC/GTA), *Msc*I (TGG/CCA) and *Ssp*I (AAT/ATT), which occur naturally in the *GUS* coding region, both the IME effect and GUS activity were abolished. To further detect the *GUS* transcript in three constructs, the mRNA level decreased to quarter of pSGN (Fig. 2). Compared with the 6-nucleotide sequence (CTGCAG) from the *Pst*I site, the three blunt-end enzyme sequences did not match the conserved border sequeces (AGgta……gcagG) of intron splicing sites, but the *Pst*I site could occur the perfect match (CTGCAGgta……ctgcagG) (Simpson and Filipowicz 1997). In prior study, Rose and Beliakoff (2000) have proved the *Pst*I was the selectable enzyme for intron splicing analysis. Thus the conserved sequences on both sides of the *ubi1* intron appeared to contribute to the correct splicing and enhanced GUS activity.

3.2. The roles of different portions of the *ubi1* intron sequence in IME

Histochemical GUS staining and fluorescence quantitation analyses of GUS activity in a series of constructs bearing deletions of the *ubi1* sequence showed different gene expression levels (Fig. 3). Deletions at the 3′ end of the *ubi1* sequence significantly reduced GUS expression; in contrast, deletions at the 5′ end of the *ubi1* sequence enhanced GUS expression to a certain point. These results indicated that both P1 and P2 sequences affected the function of IME. Plasmid pSG(13i-P3)N showed a greater enhancement than pSGN but remained lower than the construct consisting of the full-length *ubi1* intron. This evidence demonstrated that the +714 to +806 region in the *ubi1* intron is important for triggering the IME mechanism, which was further supported by internal deletions of the B sequence. Nucleotides 714 to 806 deletion alone showed a lower enhancement than the construct pSG(13i)N.

The enhancement of truncated *ubi1* intron by 5′ end upstream deletions did not clearly decrease the IME effect compared to the full-length intron. The 142-nt deletion construct (pSG(13i-P5)N) caused a 1.5-fold enhancement; however, the longer deletion (280 nt) had no advantage relative to pSG(13i-P5)N (Fig. 3-B). The other internal deletion construct pSG(13i-A)N exhibited a slightly lower enhancement than pSG(13i-P5)N but greater than pSG(13i-P4)N, which was similar to the full-length *ubi1* intron. These results suggested that nucleotides 52 to142 may inhibit the enhancement effect of the *ubi1* intron. Prior studies of *Adh1* intron 1 revealed that only long internal deletions affected steady-state RNA levels and that the artificial intron required a minimum of 70–73 nt for efficient splicing (Goodall and Filipowicz 1990; Luehrsen and Walbot 1994); these results differ from our findings. The length of the introns may lead to differences in IME efficiency between the *adh1* intron and the *ubi1* intron. The length of the first intron of the *Sh1* gene is similar to that of the *ubi1* intron. A 145-nt derivative region could cause approximately the same 20- to 50-fold stimulation, identical to the effect of the full-length *Sh1* intron (Clancy and Hannah 2002).

Many studies have revealed that the enhancing intron sequence in plants and animals contains *cis*-acting elements, including splicing enhancers and silencers (Hertel *et al*. 1997; Schaal and Maniatis 1999; Zheng 2004; Pertea *et al*. 2007; Wang and Burge 2008; Juan *et al*. 2014). These splicing regulatory elements can activate or inhibit spliceosome recognition of splice sites by recruiting *trans*-acting factors (Hertel *et al*. 1997). In the present study, the pSG(13i-P5)N construct achieved the maximum IME efficiency, and two internal deletions did not affect its performance. These results suggested that certain regulatory elements that control the initiation of IME may be in the 3′ end of the *ubi1* intron. Regulatory regions that have a negative effect on target gene expression may exist in the deleted fragment of the *ubi1* intron.

3.3. The feature of the IME signal in the artificially modified *ubi1*-P5 intron

Rose *et al*. (2008) developed a computer program called IMEter to predict the potential effect of IME; the predictions of this program were tested in *Arabidopsis thaliana* and rice (Morello *et al*. 2010). In the present study, we first used the IMEter to evaluate the mutated sequences occurring in the *ubi1*-P5 intron and found that the score positively correlated with the number of mutated motifs. Further experiments revealed that some of the intentionally mutated motifs are important for gene expression (Fig. 4-B and C). The transient transformation results showed that mutations M1 (pSG(13i-M1)N) and M2 (pSG(13i-M2)N) reduced GUS expression, that mutation M3 (pSG(13i-M3)N) caused a small increase in GUS expression, and that mutation M4 (pSG(13i-M4)N) significantly enhanced GUS expression (Fig. 4-B). We attempted to improve IME by adding additional potential enhancing motifs; however, adding one motif (pSG(13i-M1)N) or two motifs (pSG(13i-M2)N) did not cause the predicted increases. This finding suggested that simply increasing the number of motifs is unlikely to enhance the effect of IME. Sequence M4 showed the highest expression

due to its location or due to a comprehensive effect rather than an additive effect. The role of each mutation site in IME was further tested one by one; none of these sites could individually contribute to the enhancement of GUS expression (Fig. 5). These results confirmed that the *ubi1* intron adopts a complex structure that is, more inclined to an integrated form, to exert functions. The IMEter program was also verified in rice; however, the GUS expression level in transiently transformed rice calli did not significantly correlate with the calculated IMEter scores (Rose *et al*. 2008; Akua and Shaul 2013). In the present study, GUS activity increased with the number of mutated sites, which suggested that *GUS* gene expression positively correlated with the score; however, the enhancement effect was related to the number and location of the enhancing motifs in some way.

3.4. Verification of *ubi1* intron IME in rice

The pSG(13i-M4) construct was tested in bombarded rice calli to further verify the IME effect of the modified *ubi1* intron in other monocot plants (Fig. 4-D). As expected, the pSG(13i-M4) construct showed a higher enhancement than the pSGN control vector and pSG(13i)N in all of the transgenic lines. Although a significant difference in GUS expression was observed between maize and rice due to the recipient materials, the same result was observed in the two staple crops. The insertion of the *ubi1* intron into the *GUS* coding sequence could improve gene expression. In addition, because prokaryotic organisms lack a mechanism to allow intron splicing, the transgene, including the intron, would be difficult to splice and translate in the case of horizontal transfer events. This feature may alleviate concerns regarding the biosafety of developing genetically modified (GM) crops. Thus, the modified *ubi1*-M4 intron is a candidate regulatory element for elevating foreign gene expression in plant genetic engineering. It may be inserted into any target gene coding sequence after ATG or near ATG in plant expression construct, then the *ubi1*-M4 intron would be spliced from the mature RNA and the foreigner gene expression enhanced.

3.5. The border sequence of the intron-exon junction influences IME efficiency

We attempted to construct a universal vector that can easily carry any foreign genes to generate a more convenient and efficient application of the modified *ubi1* intron in further transgenic research. According to previous results, the modified *ubi1*-P5 and *ubi1*-M4 introns were inserted immediately downstream of the initiator AUG codon, and the two enzyme sites *Sal*I and *Nhe*I were introduced downstream of the intron for foreign gene insertion in the correct orientation. Three universal vectors were constructed, and GUS activity was measured in bombarded maize calli (Fig. 6-A and C). All of the universal vectors could enhance gene expression; this enhancement was defined as 2-fold higher GUS activity than the non-intronic pSGN control vector. However, inserting the intron into different positions in the *GUS* coding region caused different IME effects. The pSG(13i)N with a +13 insertion showed 4-fold higher GUS activity compared to pSAiGN, which has a +4 site insertion. Meanwhile, the vector pSG(13i-M4)N also showed the same 4-fold higher IME efficiency than the pSGAM4GN vector.

Mascarenhas *et al*. (1990) reported that the exon sequences adjacent to the intron influenced the ability of the intron to affect CAT gene expression. The exon junction complex (EJC) located 20–24 nucleotides upstream of exon-exon junctions (Le *et al*. 2000, 2001; Dimaano and Ullman 2004; Valencia *et al*. 2008) could promote the efficient export of spliced mRNA from the nucleus (Wiegand *et al*. 2003; Nott *et al*. 2004; Mufarrege *et al*. 2011), where the spliceosome deposits multiple proteins. Compared to pSG(13i)N, pSG(13i-P5)N and pSG(13i-M4)N, the three universal vectors had nine nucleotides shifted from upstream of the intron to downstream of the intron and a *Pst*I site changed to *Sal*I site (GTCGAC) in a sequence flanking the 3′ splicing site; these features may influence the IME efficiency to some extent.

To avoid this influence, three improved universal vectors were constructed in which the first 12 nucleotides (+1 to +12) of the *GUS* coding sequences were retained (Fig. 6-B). The three constructs were bombarded into maize calli and rice calli, and the level of enhancement obtained was similar to that of pSG(13i)N (Fig. 6-D). Although the enhancement effect of these improved vectors was lower than that of pSG(13i-M4)N, the IME effect was relatively less pronounced in monocots. Compared to vector pSG(13i-M4)N, the vector pSA12M4GN had an enzyme site changed from *Pst*I to *Sal*I, and its IME effect was reduced. This observation indicated that the border sequence of the *ubi1* intron is an important factor for IME efficiency.

The distance to the AUG codon differed in the three type introns: the 5′ UTR intron, CDS (the coding sequence) introns, and 3′ UTR introns; the 5′ UTR introns are preferentially close to the initiating codon (Chung *et al*. 2006). Thus, the short distance to the initiating codon is a benefit to the 5′ UTR introns; however, the initiator AUG codon acts as the translation factor binding site to promote translation in the CDS introns, and the spliceosome complex must compete with translating factors for this binding site if the 5′ splicing site is extremely close to the initiator codon. Different optimal distances to 5′ UTR introns and CDS introns may exist. Many studies have suggested that flanking sequences of an intron might contain *cis*-acting elements, including splicing

enhancers and silencers (Rose 2004; Wang and Burge 2008). In the present study, our findings confirmed that the flanking sequences do indeed influence the IME effect of *ubi1* and its derivatives.

4. Conclusion

Intron is one of the most important regulation elements in eukaryote organism. In this study, the first intron of maize ubiquitin was analyzed comprehensively including location effect, deletion analysis and mutation for adding more enhancing motif, and the result showed the enhancement of *ubi1*-P5 intron and *ubi1*-M4 intron to gene expression were higher than that of *ubi1*full length. The modified *ubi1*-M4 intron is a candidate regulation element to enhnace foreign gene expresion for the crop genetic manipulation.

Acknowledgements

This work was supported by the National Natural Science Fund of China (30970231) and the Genetically Modified Organisms Breeding Major Project of China (2014ZX08003-001).

References

Akua T, Shaul O. 2013. The *Arabidopsis thaliana* MHX gene includes an intronic element that boosts translation when localized in a 5´ UTR intron. *Journal of Experimental Botany*, **64**, 4225–4270.

Bruce W B. 1989. Photoregulation of a phytochrome gene promoter from oat transferred into rice by particle bombardment. *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 9692–9696.

Buchman A R, Berg P. 1988. Comparison of intron-dependent and intron-independent gene expression. *Molecular and Cellular Biology*, **8**, 4395–4405.

Callis J, Fromm M, Walbot V. 1987. Introns increase gene expression in cultured maize cells. *Genes & Development*, **1**, 1183–1200.

Chasin L. 2007. Searching for splicing motifs in alternative splicing in the postgenomic era. In: Blencowe B J, Graveley B R, eds., *Landes Biosciences*. Springer, Austin. pp. 85–106.

Choi T, Huang M, Gorman C, Jaenisch R. 1991. A generic intron increases gene expression in transgenic mice. *Molecular and Cellular Biology*, **11**, 3070–3074.

Chung B Y W, Simons C, Firth A E, Brown C M, Hellens R P. 2006. Effect of 5´ UTR introns on gene expression in Arabidopsis thaliana. *BMC Genomics*, **7**, 120.

Clancy M, Hannah L C. 2002. Splicing of the maize *Sh1* first introns is essential for enhancement of gene expression, and a T-rich motif increases expression without affecting splicing. *Plant Physiology*, **130**, 918–929.

Clancy M, Vasil V, Hannah L C, Vassil I K. 1994. Maize Shruken-1 intron and exon regions increase gene expression in maize protoplasts. *Plant Science*, **98**, 151–161.

Dimaano C, Ullman K S. 2004. Nucleocytoplasmic transport: integrating mRNA production and turnover with export through thenuclear pore. *Molecular and Cellular Biology*, **24**, 3069–3076.

Farquharson K L. 2008. The IMEter predicts an intron's ability to boost gene expression. *The Plant Cell*, **20**, 498.

Goodall G J, Filipowicz W. 1990. The minimum functional length of pre-mRNA intron in monocots and dicots. *Plant Molecular Biology*, **14**, 727–733.

Hertel K J, Lynch K W, Maniatis T. 1997. Common themes in the function of transcription and splicing enhancers. *Current Opinion in Cell Biology*, **9**, 350–357.

Huang M T F, Gorman C M. 1990. Intervening sequences increase efficiency of RNA 3´ processing and accumulation of cytoplasmic RNA. *Nucleic Acids Research*, **18**, 937–947.

Jefferson R A, Kavanagh T A, Bevan M W. 1987. GUS fusion: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal*, **6**, 3901–3907.

Juan W C, Roca X, Ong S T. 2014. Identification of *cis*-acting elements and splicing factors involved in the regulation of BIM pre-mRNA splicing. *PLOS ONE*, **9**, e95210.

Le Hir H, Gatfield D, Izaurralde E, Moore M J. 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO Journal*, **20**, 4987–4997.

Le Hir H, Izaurralde E, Maquat L E, Moore M J. 2000. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO Journal*, **19**, 6860–6869.

Le Hir H, Nott A, Moore M J. 2003. How introns influence and enhance eukaryotic gene expression. *Trends in Biochemical Sciences*, **28**, 215–220.

Luehrsen K R, Walbot V. 1994. Addition of A- and U- rich sequence increases the splicing efficiency of a deleted form of a maize intron. *Plant Molecular Biology*, **24**, 449–463.

Mascarenhas D, Mettler I J, Pierce D A, Lowe H W. 1990. Intron-mediated enhancement of heterologous gene expression in maize. *Plant Molecular Biology*, **15**, 913–920.

Meredith J, Storti R V. 1993. Developmental regulation of the Drosophila tropomyosin II gene in different muscles is controlled by muscle-type-specific intron enhancer elements and distal and proximal promoter control elements. *Developmental Biology*, **159**, 500–512.

Moabbi A M, Agarwal N, Kaderi B E, Ansari A. 2012. Role for gene looping in intron-mediated enhancement of transcription. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 8505–8510.

Moore M J, Proudfoot N J. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*, **136**, 688–700.

Moore M J, Sharp P A. 1993. Evidence of two active sites in the spliceosome provided by stereochemistry of pre-mRNA. *Nature*, **365**, 364–368.

Morello L, Gianı S, Troina F, Breviario D. 2010. Testing the IMEter on rice introns and other aspects of intron-mediated enhancement of gene expression. *Journal of Experimental Botany*, **62**, 533–544.

Mufarrege E F, Gonzalez D H, Curi G C. 2011. Functional interconnections of *Arabidopsis* exon junction complex proteins and genes at multiple steps of gene expression. *Journal of Experimental Botany*, **62**, 5025–5036.

Niu D K, Yang Y F. 2011. Why eukaryotic cells use introns to enhance gene expression: Splicing reduces transcription-associated mutagenesis by inhibiting topoisomerase I cutting activity. *Biology Direct*, **6**, 24.

Nott A, Le Hir H, Moore M J. 2004. Splicing enhances translation in mammalian cells: An additional function of the exon junction complex. *Genes & Development*, **18**, 210–222.

Okkema P G, Harrison S W, Plunger V, Aryana A, Fire A. 1993. Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. *Genetics*, **135**, 385–404.

Parra G, Bradnam K, Rose A B, Korf L. 2011. Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants. *Nucleic Acids Research*, **9**, 5328–5337.

Pertea M, Mount S M, Salzberg1 S L. 2007. A computational survey of candidate exonic splicing enhancer motifs in the model plant *Arabidopsis thaliana*. *BMC Bioinformatics*, **8**, 159.

Rose A B. 2004. The effect of intron location on intron-mediated enhancement of gene expression in *Arabidopsis*. *The Plant Journal*, **40**, 744–751.

Rose A B, Beliakoff J A. 2000. Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing. *Plant Physiology*, **122**, 535–542.

Rose A B, Elfersi T, Parra G, Korf L. 2008. Promoter-proximal introns in *Arabidopsis thaliana* are enriched in dispersed signals that elevate gene expression. *The Plant Cell*, **20**, 543–551.

Samadder P, Sivamani E, Lu J L, Li X G, Qu R D. 2008. Transcriptional and post-transcriptional enhancement of gene expression by the 5′ UTR intron of rice *rubi3* gene in transgenic rice cells. *Molecular Genetics and Genomics*, **279**, 429–439.

Schaal T D, Maniatis T. 1999. Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA. *Molecular and Cellular Biology*, **19**, 261–273.

Simpson G G, Filipowicz W. 1997. Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organisation of the spliceosomal machinery. *Plant Molecular Biology*, **32**, 1–4.

Vain P, Finer K R, Engler D E, Pratt R C, Finer J J. 1996. Intron-mediated enhancement of gene expression in maize (*Zea mays* L.) and bluegrass (*Poa pratensis* L.). *Plant Cell Reports*, **15**, 489–494.

Valencia P, Dias A P, Reed R. 2008. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 3386–3391.

Wang Z F, Burge C B. 2008. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA*, **14**, 802–813.

Wiegand H L, Lu S H, Cullen B R. 2003. Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 11327–11332.

Zhang X J, Liu J Q, Zhao Q, Yu J J, Ao G M. 1999. Transfer of high lysine-rich gene into maize by microprojectile bombardment and detection of transgenic plants. *Journal of Agricultural Biotechnology*, **7**, 363–367. (in Chinese)

Zheng Z M. 2004. Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *Journal of Biomedical Science*, **11**, 278–294.

(Managing editor WANG Ning)